CULC-H05881-8.271153

CALCUTTA UNIVERSITY
CENTRAL LIBRARY

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

FEBRUARY 1967

8

VOL. XXXVII, PART 1

Price One Pound

711

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by
L. B. BIRCH

Associate Editors:
P. E. VERNON and DORIS M. LEE

Advised by

SIR CYRIL BURT
R. A. C. OLIVER
E. A. PEEL
SIR FRED J. SCHONELL
S. WISEMAN

Vol. XXXVII, Part 1

The British Journal of Educational Psychology is issued by the British Psychological Society and the Association of Teachers in Colleges and Departments of Education, once a term, in February, June and November. The subscription is £2 a year post free, payable in advance, or £1 each number.

Subscriptions and Orders should be sent to "The Manager, B.J.E.P. Department, Methuen and Co., Ltd., 11, New Fetter Lane, London, E.C.4. Members of the British Psychological Society and members of the A.T.C.D.E. receive the *Journal* on special terms through the Societies.

Papers for publication should be sent to L. B. Birch, The Institute of Education, The University, Sheffield, 10. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for the *British Journal of Educational Psychology* should be sent to L. B. Birch, at the above address. *Advertisements* and *Subscriptions* for this *Journal* should be sent to Manager, B.J.E.P. Department, Sweet and Maxwell, Spon Booksellers, Ltd., 11, New Fetter Lane, London, E.C.4.

Published by
METHUEN AND CO., LTD., 11, NEW FETTER LANE, LONDON, E.C.4,
for the British Journal of Educational Psychology, Ltd.
Registered Offices: 42, Hill Street, Birmingham, 5. This address should *not* be used for letters or books intended for the Editor, Authors, or the Publishers.

On behalf of
The British Psychological Society and the Association of Teachers in Colleges and Departments of Education.

All rights reserved.

PRINTED BY THE BIRMINGHAM PRINTERS, LTD., 42, HILL STREET, BIRMINGHAM, 5.

VOLUME XXXVII FEBRUARY 1967 PART 1

FOLLOW-UP OF AN EXPERIMENT IN SELECTION FOR REMEDIAL EDUCATION*

By H. LYTTON

(*Institute of Education, University of Exeter*)

CALCU CENTRAL LIBRARY UNIVERSITY

SUMMARY. This paper reports the continuation of an earlier investigation into selection for remedial education (Lytton, 1961). The children's gains over the whole of the remedial period were recorded and each child was followed up one to two years after he left the remedial group. While children lost some of their original gains in reading as well as in arithmetic during the follow-up period, the losses in arithmetic were much greater. Judged by their standing in relation to their age groups (Achievement Ratio), some children maintained a large part of their gains after remedial teaching ceased, especially in reading.

In reading, selection by tests and selection by teachers' judgment did not produce significantly different long-term results. In arithmetic, higher gains were associated with test selection. Both in reading and arithmetic, children with high I.Qs. recorded greater gains during the remedial period and maintained them better afterwards than children with a low I.Q. However, in reading—as opposed to arithmetic—even children of below average intelligence maintained some of their gains. A.Q. was an unsatisfactory predictor of success. Those who had changed to secondary schools during the follow-up period lost more than those who were still in primary schools. Maladjustment was found to be inversely related to long-term gains due to remedial education.

I.—INTRODUCTION.

IN 1961 the author reported the first year's results of an experiment in remedial education (Lytton, 1961). The present paper reports the results for the total remedial period and for a follow-up.

Table 1 provides some basic data about the children.

TABLE 1

BASIC DATA ON SUBJECTS.

		Reading			Arithmetic		
		Test-selected children	Teacher-selected children	Overall	Test-selected children	Teacher-selected children	Overall
N		41	56	97	17	23	40
I.Q.	M S.D.	96·05 —	87·52 —	91·12 14·53	100·47 —	85·13 —	91·65 13·49
Init. A.Q.........	M S.D.	74·68 —	83·63 —	79·85 14·99	80·77 —	89·39 —	85·73 12·52
Init. A.R.	M	71·68	71·50	71·58	81·53	76·83	78·83
Init. C.A.	M	8–10	9–2	9–0	9–2	9–5	9–3
Init. R.A. (A.A.)	M	6–3	6–6	6–5	7–5	7–2	7–3
Months of remed. period	M	18·28	14·37	16·15	6·53	12·14	9·78

---

* This paper is based on part of a thesis submitted to the University of London for the degree of Ph.D.

II.—THE PRESENT INVESTIGATION.

After an average interval from the end of remedial teaching of 16·4 months in the case of reading and 18·5 for arithmetic a follow-up was carried out on those children who were still in the schools of the local authority. Their numbers had dropped from ninety-seven to eighty-eight for the reading group and from forty to thirty-seven for the arithmetic group. Reading was assessed by the Burt-Vernon Graded Word Reading Test (Vernon 1938), by a comprehension test (Schonell 1950) and by a grading by the writer for accuracy, fluency and understanding after hearing each child read the book he was using in class. In the case of arithmetic the child was again given the Staffordshire Arithmetic Test, which had been used previously, and the written arithmetic he did in class was scrutinised for competence and any special difficulty.

At the same time, the child's class teacher filled in a " Report Form " which consisted mainly of questions on the child's behaviour, designed to elicit indications of maladjustment. It was used by Bowlby and his co-workers (1956) in a study of the effects of mother-child separation as observed in children in a sanatorium. The teacher was also asked to rate the child's attainments in reading and/or arithmetic on a four-point scale. When the teacher's and psychologist's ratings of reading diverged widely, it was usually possible, after a personal discussion, to illuminate the reasons for the discrepancy and to narrow it. The two ratings were then combined. In arithmetic the teacher's rating alone was used.

The Gain per month (Follow-up), the Comprehension A.R. (R.A./C.A.) and the teacher's and psychologist's ratings, all scaled on a common scale, were added to form the 'Combined Criterion' for reading. For arithmetic, the 'Combined Criterion' is composed of the Gain per month (Follow-up), the Arithmetic A.R. at follow-up and teacher's assessment of arithmetic, all scaled on a common scale.

III.—RESULTS AND DISCUSSION.

The results are summarised in Tables 2, 3, 4, 5 and 6. The analysis was carried out by a formula that allowed for systematic differences between adjustment groups, suggested by Dr. D. N. Lawley, University of Edinburgh.*

TABLE 2

CRITERION SCORES—MEANS, S.D'S, RANGES OVER ALL SCHOOLS.

		Gain per month of remedial period	Gain per month of follow-up period	Comprehension AR	Combined Criterion
Reading	M	2·21	0·22	76·14	25·76
	S.D.....	·95	·72	10·67	6·49
	Range ..	0·1 to 4·4	–1·83 to +1·59	55 to 98	12 to 38
Arithmetic	M	2·74	·00	—	16·08
	S.D. ...	1·90	·90	—	7·26
	Range ..	0·0 to 8·5	– 2·58 to +2·29	—	5 to 33

---

* The comparisons between the various groupings shown in the tables were based on the simple averages (unweighted means) of the mean scores of the separate schools. The variance was computed by pooling the within-schools sums of squares and the variance of the mean of three schools then obtained by the formula :

$$\sigma^2_m = \frac{\sigma^2}{3^2}\left(\frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3}\right)$$

where σ^2 = pooled variance and n_1 = number of subjects in school 1, etc.
The t-test was then applied in the usual way.

An analysis by sex showed boys and girls to have virtually identical criterion scores. Boys' and girls' results were, therefore, combined for all the analyses shown.

(1) *Reading.*

(*a*) *Teacher v. Test Selection.* In spite of the higher I.Q. of the test-selected children, there is no difference in criterion scores of achievement and this confirms the results after the first session's teaching.

TABLE 3

READING RESULTS.

(I.Qs. based on Sleight Non-Verbal Intelligence Tests)
Unweighted Means of three Schools.

Groupings	N[1]	Gains per month of remedial period	Gains per month of follow-up period	Comprehension AR	Combined[2] Criterion
Test-selected	41	2·27	0·22	77·30	26·17
Teacher-selected ..	56	2·21	0·21	75·17	25·39
I.Q.					
(1) 63–84	34	1·96 ←	0·06	69·84 ←← ←←←	22·65 ←←←
(2) 85–99	33	2·26	0·24	76·96 ←← ←	25·57 ←
(3) 100+	30	2·50 ←	0·39	82·71 ← ←←←	29·37 ← ←←←
Initial A.Q.					
(1) 46–80	56	2·18	—	—	24·35
(2) 81+..........	41	2·27	—	—	27·16
Initial A.R.					
(1)<80	76	2·20	0·15[3]	74·47[3]	24·67 ←←
(2) 80+..........	21	2·25	0·50	84·00	30·29 ←←
Initial R.A.					
(1) 5·0—5·11	40	2·09 ←	—	74·24[3]	23·76 ←
(2) 6·0—6·11	31	2·15	—	75·84	25·77
(3) 7·0+	26	2·62 ←	—	78·01	27·64 ←
G.p.m. (Remed. Period).					
(1) 0—1·5	19[1]	—	0·06	—	—
(2) 1·6—2·5	36	—	0·10 ←←	—	—
(3) 2·6+	33	—	0·53 ←←	—	—
Primary School Follow-up	52[1]	2·36	0·36 ←	80·63 ←←←	27·72 ←←
Secondary School Follow-up	36	2·24	0·02 ←	69·81 ←←←	23·11 ←←

[1] Some subjects were lost before Follow-up.

[2] Composed of
(*a*) Gain per month (Follow-up).
(*b*) Comprehension AR.
(*c*) Teachers' and psychologist's ratings of reading.

[3] Not tested because difference in same direction as Comb. Crit.

NOTE:
Single arrow indicates difference significant at 5 per cent. level.
Double arrow indicates difference significant at 1 per cent. level.
Treble arrow indicates difference significant at 0·1 per cent. level.

(*b*) *I.Q.* The difference in mean gain during the remedial period between the lowest and the highest I.Q. groups is significant. (Differences for the long-term results generally are more marked than in the first session alone).

(*c*) *Initial Achievement Quotient* (*A.Q.*= *R.A.*/*M.A.*). A 2×3 analysis of variance of Combined Criterion scores was carried out for I.Q. and A.Q. groupings (Table 4). This is statistically not entirely sound, because differences between adjustment groups could not be allowed for, but it may be considered as indicating suggestive trends. I.Q. and A.Q. differences were each significant by themselves. The interaction I.Q./A.Q. was also significant. At a low I.Q. level children in the high A.Q. group did very much better than children in the low A.Q. group and this difference progressively decreased at each higher I.Q. level.

TABLE 4

READING—COMBINED CRITERION SCORES ANALYSED BY I.Q. AND A.Q. GROUPS.

I.Q.	A.Q.		
	48—80	81+	Overall
63—84	17·55	25·40	22·61
85—99	23·00	29·67	25·67
100+	29·00	31·60	29·48
Overall	24·11	27·62	25·76

(*d*) *Initial Achievement Ratio* (*A.R.* = *R.A.*/*C.A.*). Gains during the remedial period were not materially different for the low and the high initial A.R. groups (nor had they been in the first session results).

At follow-up a highly significant difference appeared between the good ultimate adjustment of the original high A.R. group and the lesser standing of the lower A.R. group, as seen in the Combined Criterion. Differences in Comprehension A.R. and in G.p.m. (F) were not tested, as the picture seemed much the same.

(*e*) *Initial R.A.* The higher the initial R.A. the greater the gain in the remedial period, and the greater the ultimate adjustment of the child, as seen in the significant difference between Combined Criterion scores. Differences in Comprehension A.R. and G.p.m. (F) were again not tested, as the situation seemed to be the same as for the Combined Criterion. (In the first session differences were extremely small and insignificant.)

(*f*) *G.p.m. Remedial Period.* It is interesting to note that differences in gains during the remedial period were on the whole, reflected accurately by by differences in the follow-up period.

(*g*) *Primary/Secondary School Follow-up.* All the remedial teaching was done at primary schools. It was thought that whether a child afterwards transferred to a secondary school or stayed in a primary school might affect the maintenance of his gain and the results were, therefore, analysed according to whether the child was in a primary or secondary school at follow-up.

Indeed, a small and insignificant difference in favour of the 'Primary School Follow-up' children at the end of the remedial period turned into large and highly significant differences at follow-up. It must be pointed out, however, that the mean I.Q. of the ' Primary School Follow-up ' children was 92·9 and of the ' Secondary School Follow-up ' children 87·4.

(2) *Arithmetic.*

TABLE 5

ARITHMETIC RESULTS

(I.Qs. based on Sleight Non-Verbal Intelligence Tests)
Unweighted Means of Two Schools.

Groupings	N[1]	Gains per month of remedial period	Gains per month of follow-up period	Combined[2] Criterion
Test-selected	17	3·41	0·51 ←←	20·56 ←←
Teacher-selected	23	2·35	–0·32 ←←	13·50 ←←
I.Q.				
(1) 63–84	13	1·58 ← ←	–0·52 ←	10·36 ←←←
(2) 85–99	15	3·38 ←	0·35 ←	19·44 ←←←
(3) 100+	12	3·12 ←	0·22	18·88
Initial A.Q.				
(1) 56–84	15	3·07	0·02	16·17
(2) 85+	25	2·67	0·03	16·49
Initial A.R.				
(1) <80	16	3·22	–0·01	16·17
(2) 80+	24	2·73	0·16	17·46
Initial A.A.				
(1) 6·0—6·11	21	2·23	—	13·37 ←
(2) 7·0+	19	3·17	—	19·28 ←
G.p m. (Remed. Period)				
(1) 0—1·5	12[1]	—	0·06	—
(2) 1·6—2·5	8	—	–0·60	—
(3) 2·6+	17	—	0·17	—

[1] Some subjects were lost before Follow-up.

[2] Composed of : (*a*) Gain per month (Follow-up).
(*b*) Follow-up A.R.
(*c*) Teacher's assessment of arithmetic.

NOTES :
Single arrow indicates difference significant at 5 per cent. level.
Double arrow indicates difference significant at 1 per cent. level.
Treble arrow indicates difference significant at 0·1 per cent. level.

(*a*) *Teacher v. Test Selection.* The difference in mean gain was not significant in the first session, it approached significance at the end of the remedial period and differences became very significant at follow-up.

(*b*) *I.Q.* It is notable that in all criterion measures the I.Q. group 85–99 did best in the first session, at the end of the remedial period, and at follow-up. The differences between the two higher groups remained insignificant, but those between the lowest group and both higher ones always reached at least the ·05 level, the difference in the Combined Criterion being the most significant.

(*c*) *Initial Achievement Quotient* (*A.Q.* = *A.A./M.A.*). A 2×3 analysis of variance of G.p.m. (R) showed a significant I.Q./A.Q. interaction. The higher A.Q. group did very much better at the middle level of I.Q., than at either the low or the high level of I.Q. (Table 6).

TABLE 6

ARITHMETIC—G.P.M. (R) ANALYSED BY I.Q. AND A.Q. GROUPS.

I.Q.	A.Q.		
	56—84	85+	Overall
63—84	2·10	1·66	1·72
85—99	1·53	3·76	3·31
100+	3·53	1·05	3·12
Overall	2·94	2·62	2·74

An analysis of variance of G.p.m. (F) did not show a significant I.Q./A.Q. interaction.

(*d*) *Initial Arithmetic Achievement Ratio* (*A.R.*=*A.A./C.A.*). At the end of the remedial period, as in the first session, the lower A.R. group had made more progress, but at follow-up this was reversed. No difference, however, was significant.

(*e*) *Initial A.A.* Those with a higher initial arithmetic age made slightly more progress during the remedial period and, above, all, were of significantly better standing at follow-up.

(*f*) *G.p.m. Remedial Period.* There is no clear relationship in arithmetic, as opposed to reading, between gain during the remedial period and gain during the follow-up period.

(*g*) *Primary/Secondary School Follow-up.* The results are not included in the table, as they are too contradictory to be meaningful, the contradiction being due mainly to exceptional scores of small sub-groups.

(3) *Achievement in relation to age.*

TABLE 7

AVERAGE ACHIEVEMENT RATIOS (R.A./C.A.).

	Overall	Test-select.	Teacher-select.	I.Q. Below 85	I.Q. 85–99	I.Q. 100+
		READING				
Initial A.R.	71·6	71·7	71·5	68·9	71·4	74·9
Final A.R.	86·6	88·6	85·2	83·4	87·9	89·4
F'up A.R.	80·8	82·2	79·6	75·2	81·7	86·0
		ARITHMETIC				
Initial A.R.	79·4	81·6	78·1	76·6	80·9	81·3
Final A.R.	89·7	92·8	87·5	83·8	92·8	91·6
F'up A.R.	79·3	87·0	74·0	67·9	84·9	84·5

Table 7 shows the achievement ratios (R.A./C.A. or A.A./C.A.) for the beginning and end of the remedial period and for the follow-up. There is, typically, a gain during the intensive remedial period, i.e., the children move markedly towards the 'normal' achievement ratio of 100, and, equally typically, a loss afterwards.

In reading, differences between the different I.Q. groups are very marked, but even the lowest I.Q. group (63–84) retains some of its gains. In arithmetic those with I.Qs. above 85, retain a considerable part of their gains, those below that line deteriorate. Selection by tests rather than by teachers' judgment seems to be effective here.

(4) *Maladjustment Index.*

From the Teacher's Report Form, the number of items indicating maladjustment was computed to give a 'Maladjustment Index.' Rank correlations showed a highly significant association between absence of maladjustment (Maladjustment Index reversed) and Combined Criterion, both for reading and arithmetic, but immediate progress in the remedial group (gain per month of remedial period) seemed to have a much more tenuous association, with absence of maladjustment for reading and none at all for arithmetic (Table 8).

TABLE 8

RANK CORRELATIONS BETWEEN ABSENCE OF MALADJUSTMENT, COMBINED CRITERION AND G.P.M. (R).

Maladjustment Index reversed *v.*	Rho	P (1-tailed test)
Combined Criterion		
Reading	·26	·01
Arithmetic	·47	·005
Gain per month of remedial period :		
Reading	·16	·10
Arithmetic	—·05	N.S

IV.—CONCLUSIONS.

The numbers involved in this investigation are as large as they normally are in experiments in remedial education. Nevertheless, looked at from the point of view of stability of results they are rather small and conclusions based on them must accordingly be treated with reserve until they have been confirmed by a repeat of the experiment.

(1) In reading there is no significant difference in effectiveness between selection by tests and selection by teachers' judgment. Selection by teachers' judgment, which is simpler and will give opportunities to some children who are more in need of remedial education, but of lower I.Q., can be used without fear of loss of 'profitability.'

In arithmetic higher gains are definitely associated with selection by tests, and this may, therefore, be preferred. In this respect, first session conclusions are revised.

(2) In both reading and arithmetic, the I.Q.—even though derived from Picture Tests—has proved a good pointer to relative success in remedial work. In reading, even the children of below average intelligence make worthwhile

progress and seem to maintain some of it over the follow-up period. In arithmetic, however, the children in the lowest I.Q. group seem to derive no lasting benefit in terms of attainments from the special tuition, and their inclusion in remedial groups, such as described here must, therefore, be debatable. Here, too, first session conclusions have been modified through the later results, and the tentative nature of the conclusions should be especially emphasised.

(3) The insignificant differences between the A.Q. groupings in reading and arithmetic show this measure to be an unsatisfactory predictor, and this is in line with its instability expected on theoretical grounds (cf. Vernon 1958, Crane 1959). The results for reading seem to show that within the lowest I.Q. group children of high A.Q. succeed much better than those of low A.Q.

The A.Q. has, in the final analysis, turned out to be a much less good predictor than it seemed on the basis of first session results alone.

(4) Further pointers to success were: in reading, a high initial A.R. (R.A./C.A.) and a high initial R.A.; in arithmetic, high initial A.A. (None of these measures had any prognostic significance in the first session.)

(5) In reading, the children followed up in the secondary school had lost far more of their gains than those followed up in the primary school. The difference in I.Q. between these two groups alone does not appear sufficient to explain the greater loss. (There is a larger I.Q. difference between the test-selected and the teacher-selected children without a parallel difference in long-term gain.) It is probable, therefore, that the teaching in the secondary school, being less geared to the basic skills and these children's difficulties, contributed to the relative decline in their attainments.

(6) If we take the 'maladjustment index' of this remedial group as a whole, and compare it with that of the 'normal' control group (cf. Bowlby, *et al.* (1956), or if we take the rank correlation between Maladjustment Index (reversed) and Combined Criterion, it is clear that both backwardness and small long-term gain from remedial teaching are related to maladjustment. This emerges even more strikingly when the Maladjustment Index of the very best achievers is compared with that of the very poorest.

A study of individual profiles as part of the larger investigation will, it is hoped, throw further light on this question.

(7) How far the various categories of teachers' ratings are effective in predicting success in remedial treatment forms the subject of a separate but related study. However, that selection by teachers' judgment is an effective procedure, at least for reading, has been confirmed by this follow-up. Provided it is confined to remedial reading, there is nothing here that would require the writer to change the conclusion expressed in the earlier paper: "So long as general guiding principles are borne in mind, such an informal method (of selection) would seem a sensible approach."

This study was not designed to provide direct evidence on the effectiveness of remedial teaching in producing permanent improvement and, therefore, did not include a control group of children not receiving remedial education. However, as shown above, it is clear that *some* children benefit considerably and also maintain a large part of their gains after remedial teaching has ceased, especially in reading. (58 per cent. of children had reached a reading age of 9·0+ at follow-up, a worth-while achievement for them individually.) The attempt to discover objectively whether the children concerned benefited in attitude to learning and general adjustment, as described in the earlier paper, unfortunately bore no fruit. However, if long-term gain in attainments is the main object of remedial education, it is no doubt possible to maximise it by judicious selection.

V.—REFERENCES.

BOWLBY, J., *et al.* (1956). Effects of mother-child separation: a follow-up study. *Brit. J. Med. Psychol.*, 29, 211-247.

CRANE, A. R. (1959). An historcial and critical account of the accomplishment quotient idea. *Brit J. Educ. Psychol.*, 29, 252-259.

LYTTON, H. (1961). En experiment in selection for remedial education. *Brit. J. Educ. Psychol.*, 31, 79-94.

SCHONELL, F. (1950). *Diagnostic and Attainment Testing.* Oliver and Boyd.

VERNON, P. E. (1938). *The Standardisation of a Graded Word Reading Test.* Publication of the Scottish Council for Research in Education, No. XII, U.L.P.

VERNON, P. E. (1958). The relation of intelligence to educational backwardness. *Educ. Rev.*, 11, 7-15.

(*Manuscript received 13th May*, 1965)

FAMILY ENVIRONMENT AND 11+ SUCCESS: SOME BASIC PREDICTORS

By D. F. SWIFT
(*University of Oxford, Department of Education*)

SUMMARY. Our ability to analyse the socio-cultural environment is vital to the development of research into the effects of social experience upon academic performance. This paper argues that, in the course of such development, there is an inevitable tendency to treat simple statistical predictors as indicators of causal mechanisms. An example of this problem is found in the way in which social class is often treated as a single experimental variable. In the present state of research sophistication it is better to treat social classes as social contexts within which particular factors may have special effects upon cognition and motivation. Analysis of sample data suggests that the effects of certain environmental variables are different in the different social classes.

I.—INTRODUCTION: CORRELATIONAL DATA IN SOCIAL RESEARCH.

SOCIO-CULTURAL analysis is beginning to enter a phase which promises sound information on the mechanisms by which social experience influences academic motivation and performance. However, in Britain, there are still some methodological misapprehensions which ought to be cleared up. These stem principally from an attempt to apply the logic of laboratory experiment to analysis of social stimulation. Reasonable though the attempt may be, it has led to a number of implicit assumptions in research which have encouraged the persistence of invalid ' knowledge.' In laboratory research, it is easy to assume, or ensure, that all potentially contaminating variables have been controlled or randomised. This is not possible in outside research and simple correlation techniques become less meaningful if only because of the much greater risk that they will uncover spurious associations.

This paper reports some findings which have been arrived at in conventional fashion by relating environmental factors to performance, and the intention will be to highlight the dangers of assuming that any significantly related factor is also a causal one. In order to do this social class has been treated as a major confounding factor in any relationship. The need for such an approach has been brought home to us by a great deal of recent research and writing. Unfortunately, the level of sophistication at which social class is dealt with in research is much lower than the theoretical conceptualisation of the problem. Our present method of defining a social class upon occupational criteria is a first step toward *categorising* the social environment rather than an efficient operational definition of a single variable. In this paper, therefore, the intention is to treat social class as a set of environments within which any particular variable may have a special relationship with performance. The different social classes provide different contexts for any single variable and so may be a statistical confounding factor which encourages us to believe that we have found an aspect of the environment which actually affects performance in school when we have not.

The scientific study of social factors in educational achievement dates at least from Binet who was well aware that his tests were tests of learned experience. Since that time there has been an accumulation of knowledge about the factors in the environment which relate to school performance. We know that children of well-educated parents will do better than children of the less well-educated; that the richer the parent the more academically able the child; and so on. (Barber, 1961; Douglas, 1964; Elder, 1965; Fraser, 1959; Halsey, 1961; Rogoff, 1961.) However, the developing sciences of social anthropology and sociology are becoming better able to assist the psychologist in appreciating the weaknesses of the ex *post facto* reasoning which has tended to follow the discovery of these gross relationships.* This reasoning has tended to assume that any 'factors' discovered in research are causal. But, as Thurstone once put it: "a correlation is a confession of ignorance." (Eysenck, 1964, p. 7). Instead of treating the factors as causal agents, they should be seen as useful correlates of particular social environments yet to be described in terms which elucidate the actual mechanisms by which aspiration and ability are influenced. The search for these mechanisms is not a purely pragmatic one in the present state of our knowledge and it is vital that findings should be related to an adequate socio-cultural conceptual framework. Otherwise, those factors which are found to relate to performance cannot be seen as anything more than simple predictors of ability and aspiration (Herriott, 1963; Swift, 1966).

II.—DESIGN OF THE RESEARCH.

The study was concerned with the relationship between family environment and the 11+ performance. A 10 per cent. sample of two local education authority divisions yielded forty-five 'middle class' and eighty-seven 'working class' families whose children were to take the 11+ in a particular year. This sample was found to reproduce the major relevant characteristics of the universe from which it was drawn. (Swift, 1965). Five aspects of the family environment are to be reported in this paper: economic characteristics, family structure, occupational status, parental educational experience and attitudes. All of these have been shown to operate as predictors of achievement and consequently all have been thought to have a causal effect upon it. When this assumption is made, it implies not only that factor X is a causal one, but that it is causal in the same way in all situations. This dangerous assumption can obscure mechanisms which might be important in certain socio-cultural environments but unimportant in others. Several of the predictors to be treated in this article will be seen to be of this nature.

III.—RESULTS.

Economic Factors and Success. Economic factors are always a first consideration in studies of this sort. As a first step in that direction, a rudimentary index of material prosperity was constructed for each family. A mark was given for each of the following: House ownership; Car ownership; Father's weekly take-home pay more than £13; Rooms-persons ratio of the family greater than one. Table 1 records the results.

From the first sub-table it can be seen that the index is significantly associated with success ($\phi = \cdot 35$, $P < 0\cdot 001$), the distribution of successful scores having a higher mean than that for the unsuccessful scores. This raises an important question. Is it a genuine and direct relationship, requiring psychological, sociological or genetic explanation, or is it a purely statistical one which has been produced by the presence of an intervening variable—social class?

---

* It is also perfectly possible, in this area, for ex *post facto* reasoning to take place *before* the study has been carried out or written up. The probability that research will find certain of these relationships gives great scope for specious theorising before the data are collected.

TABLE 1

The Relationship of the Material Index with Success or Failure in the 11+

	Material Index Score					N	Mean Score
	0	1	2	3	4		
All Parents :							
Success	0	8	9	12	8	37	2·54
Failure	27	30	17	9	12	95	1·46
Total	27	38	26	21	20	132	1·77
	$\chi^2=15\cdot69$ $P<0\cdot001$* (0, 1 *v.* 2–4) ϕ† = ·35						
Middle Class :							
Success	0	3	5	11	8	27	2·89
Failure	0	2	3	2	11	18	3·22
Total	0	5	8	13	19	45	3·00
	Not significant.						
Working Class :							
Success	0	5	4	1	0	10	1·60
Failure	27	28	14	7	1	77	1·00
Total	27	33	18	8	1	87	1·11
	P‡ = 0·155 (0, 1 v. 2–4).						
All Parents :							
Middle Class	0	5	8	13	19	45	3·02
Working Class	27	33	18	8	1	87	1·15
Total	27	38	26	21	20	132	1·76
	$\chi^2=39\cdot71$ $P<0\cdot001$ (0, 1 v. 2–4) ϕ = ·55.						

* Throughout this paper P has been found for χ^2 with one degree of freedom.

† ϕ is a coefficient of association which may be found in 2 × 2 tables either by applying a product-moment formula or by 'norming' the χ^2 statistic. In these tables it has been computed from χ^2 according to the formula $\phi = \sqrt{\frac{\chi^2}{N}}$. It has been used because in these cases it is numerically equivalent to Pearson's r.

‡ Where no χ^2 value has been given, P has been calculated by Fisher's exact test because an expected cell value is less than five.

The second sub-table answers this question as far as the middle class is concerned. The unsuccessful families actually have a higher mean score than the successful. For the working class, whilst a relationship appears to exist, the probability of its occurring by chance is too great. Despite this fairly large difference in the two means, we must conclude that the very strong connection between the material index and success is due to an even more powerful relationship between the index and social class ($\phi = \cdot55$, $P < 0\cdot001$).

Fraser (1959) correlated the per-person income of the family with I.Q., and with an attainment score, both of which were used in the secondary school selection procedure at Aberdeen. There, the correlation between Income and I.Q. was found to be ·35 and between Income and a measure of secondary school performance, ·44. In the Lancashire study, the figures for the Material Index were of a similar order :

Material Index and V.R. Score	$r = \cdot 44$
Material Index and English Score	$r = \cdot 45$
Material Index and Arithmetic Score	$r = \cdot 38$
Material Index and 11+ Success	$\phi = \cdot 35$*

Floud (1956) found that, in Middlesborough, there was a different success rate in each social class if the families were grouped according to parental income or type of housing. On the other hand, there was no such difference in South-West Hertfordshire. The implication for the authors was that, general living standards in Middlesborough had not reached a minimum level at which the effect of ' poverty ' peters out. All people appeared to have passed this point in the. much more prosperous Hertfordshire and Floud suggested that any differences in performance between the classes had to be explained by much more subtle ' social factors.'

As far as material wealth is concerned, the conclusions of this study are the same as those drawn for South-West Hertfordshire—that lack of wealth, of itself, is not an influential factor in success on the ' 11+ ' procedure. When the different items in the material index were individually compared with 11+ success, two interesting associations emerged which support this conclusion. These were interesting in that the objective facts seem to be acting as indicators of certain attitudes, rather than working themselves to produce higher scores. For example, home ownership on the part of the working class was significantly related to success. It would be hard to make out a case in support of the idea that this relationship came about through the beneficent influence of better housing conditions upon the intellectual functioning of the children. Frequently the owned houses were inferior to a good council house—old terraced houses, a cottage and some small semi-detached houses could not be put in a category higher than the general run of working class houses. The explanation must be sought in the clue to parental ideology which ownership provides. Similarly, in the middle class, a tendency *against* the ownership of cars on the part of parents of successful children and the higher mean score on the material index for parents whose children failed the 11+, provide strong support for the idea that social perceptions and values are of prime importance.

Brothers and Sisters. A great deal of reaserch has suggested that the size of a child's family associates with his ability to score in tests and most impressive evidence is offered by Nisbet (1953) to this effect. Fraser found correlations in the Aberdeen study of —·40 between Family size and I.Q., and —·46 between Family size and secondary school performance. It was expected that the sample understudy would show the same relationship, but will be seen from Table 2 that this did not happen.

There was generally a relationship in the working class between family size and ordinal position of the child on the one hand, and lack of ability to score on both tests on the other. In the middle class, there was no correlation significantly greater than zero. No correlations were as strong as those which Fraser found and further analysis of the data revealed that the first-born child has no advantage over the remainder in any social class.

---

* ϕ is not identical to r calculated on undichotomised distributions.

TABLE 2

Sample: CORRELATION BETWEEN 11+ TEST SCORES AND FAMILY SIZE.

	Verbal Reasoning		Arithmetic	
	No. in Family	Ordinal Position	No. in Family	Ordinal Position
All Working	—·207*	—·337**	—·143†	—·190‡
All Middle	+·09*	+·06	+·235	+·138
Unsuccessful Middle	—·361†	—·277	—·294	—·207
Successful Middle	—·239	—·004	+·180	+·215
Unsuccessful Working	—·075	—·135	—·035	—·125
Successful Working	—·242	—·066	—·352	—·485*

†—Approaching significance at $P>0\cdot10$.
‡=Significant with $P<0\cdot10$.
*=Significant with $P<0\cdot05$.
**=Significant with $P<0\cdot01$.
All other correlations were not signficant.

An explanation for this contradiction of earlier findings may lie in the greater numbers which were included in the Fraser sample, but a second one might be the presence of large numbers of Roman Catholics in the writer's sample. Floud (1956) found that there was no tendency amongst Catholic families for the members of the larger families to have lower I.Qs., and the same appears to be true in this sample. Catholics are particularly evident in the successful middle class group and account for the unexpected positive tendencies of the correlations. In the small numbers which are involved here, this tendency for a large number of Catholics to enter the sample as both middle class and successful families may or may not be representative of the situation in the universe. It is true to say, however, that middle class Catholics had a high success rate in this sample.

Apart from sheer numbers of siblings the sort of secondary school which they attend is usually thought to have some connection with the child's performance. This can be expected in two ways. Firstly, as a cause—because older siblings are at the grammar school, it is likely that the child will be encouraged to try to follow them. Secondly, the relationship can be expected as an effect—the social environment which produced one '11+' success should tend to produce another.

Table 3 describes how older siblings' secondary school did associate with success.

Taking the sample as a whole, the relationship was extremely clear remaining at an acceptable level of significance in the middle class, but at an unacceptable one in the working class. Since the association between success on the 11+ and an older sibling in grammar school ($\phi = \cdot414$) is stronger than that between middle class membership and an older sibling in the grammar school ($\phi = \cdot366$) the conclusion is that a relationship exists between success and grammar school siblings which is partly independent of social class membership.

Parental Occupational Status. We know that father's occupational status is a predictor of 11+ success and failure. In addition, mother's job before marriage is always assumed to have great importance for success in school and

TABLE 3

THE RELATIONSHIPS BETWEEN HAVING AN OLDER SIBLING IN THE GRAMMAR SCHOOL AN SUCCESS IN THE 11+ AND MEMBERSHIP OF THE MIDDLE CLASS

Older sibling in the grammar school associated with :			
11+ success			Membership of Middle Class
Within Total Sample	Within Middle Class	Within Working Class	Within Total Sample
$\chi^2=13{\cdot}21$	—	—	$\chi^2=10{\cdot}29$
$P<0{\cdot}001$	$P=0{\cdot}07$	$P=0{\cdot}15$	$P<0{\cdot}01$
$\phi=\cdot414$	—	—	$\phi=\cdot366$
$N=77$	$N=24$	$N=53$	$N=77$

the general opinion is that where working class children are successful, a middle class mother is one of the most likely causes. For this reason, mother's previous occupation was tabulated against success with the results summarised in Table 4.

TABLE 4

THE RELATIONSHIP BETWEEN MOTHER'S PREVIOUS OCCUPATION AND SUCCESS IN THE 11+ AND MEMBERSHIP OF THE MIDDLE CLASS.

Mother's previous occupation associated with :			
11+ success of child			Middle Class Mother married to a Middle Class Husband
Middle Class Mothers within Total Sample	Working Class Mothers within Middle Class	Middle Class Mothers within Working Class	Within total sample
$\chi^2=10{\cdot}15$	$\chi^2=4{\cdot}99$	—	$\chi^2=33{\cdot}33$
$P<0{\cdot}01$	$P<0{\cdot}05$	Independent	$P<0{\cdot}001$
$\phi=\cdot294$	$\phi=\cdot333$	—	$\phi=\cdot502$
$N=132$	$N=45$	—	$N=133$

In this sample, the relationship is significant but much less strong than that between husband's occupation and success. A fact to be noted first is that, due to the coding scheme, which included shop assistants in the routine clerical grades, there is a much higher proportion of original middle class wives in the sample than middle class husbands. This was not important to the father's group, but raised the middle class proportion to over 50 per cent. amongst wives. It is not surprising, therefore, that the relationship with success should turn out to be the mirror image of the expected picture. Consequently, in the third column of Table 4 we see that middle class wives bore no association with success amongst working class families. On the other hand, working class wives had a very definite negative association with 11+ success amongst middle class families.

The next step was to consider the importance of similar social class membership of both parents. The fourth column shows that when both parents have middle class occupations, they have a very great probability of being successful. On the other hand, when both parents have working-class occupations, there is a powerful association with failure. Between these two extremes, the middle class husband who marries a working class wife appears to be taking a most irresponsible risk with the future prospects of his children, for, in taking this step, he lowers their chances of 11+ success from 75 per cent. to 35 per cent. !

These tables certainly seem to confirm the commonly held idea that when a child performs in a way that would not be expected from the social class position of his father, some explanation may be found in the previous occupation of his mother. In this sample, a working-class wife is an important aspect of failure. Usually, this conclusion is put the other way round—that successful working class children tend to have middle-class mothers. The evidence for this is less strong.

Original Parental Occupational Status. Considering the social class of parents in terms of their present occupations allows room for a step in mobility before their occupational level is reached. In other words, people who are rated as middle class according to occupation, may have achieved it from a working class background, perhaps even during the lifetime of the '11+' candidate. Data on the social class origins of parents pushes further back into the social experience which may affect their present ideologies and perceptions, and ultimately influencing the child's ability to perform.

TABLE 5

THE RELATIONSHIP BETWEEN GRANDFATHERS' SOCIAL CLASS AND SUCCESS ON THE 11+.

Middle Class Membership of Grandfathers associated with 11+ Success of Child.		
Within Total Group of Grandfathers	Within Paternal Grandfathers	Within Maternal Grandfathers
$\chi^2=8{\cdot}49$	$\chi^2=9{\cdot}70$	$\chi^2=1{\cdot}364$
$P<0{\cdot}01$	$P<0{\cdot}01$	Not significant.
$\phi={\cdot}179$	$\phi={\cdot}271$	—
N=264	N=132	N=132

In Table 5 we see that parental middle class origins relate with success. On differentiating between maternal and paternal grandparents, a quite strong difference emerges. Father's original social class repeats the relationship ($\phi={\cdot}271$) whereas mother's original social class is not significantly related. This is an interesting point which has some bearing upon the genetic explanation for the differing ability levels in the different social classes. Assuming that both sexes make equal contributions to the genetic potential of the child, one would expect maternal and paternal contributions to be similar.

Parents' Education. Inevitably, the educational experience of the parents will affect the attitudes towards the school which the child develops. Better 'educated' people will have a way of looking at the world and solving its problems which are likely to make a difference to the child's performance at school. How this affects selection within the school system is an important consideration in a study of the links between social background and educational performance.

TABLE 6

The Relationship Between Parental School Type and Success on the 11+.

Parents' Grammar School Associated with Child's Success on 11+.				
Within Total Sample Mothers and Fathers	Within Middle Class		Within Working Class	
	Husbands	Wives	Husbands	Wives
$\chi^2=47{\cdot}49$	$\chi^2=1{\cdot}20$	$\chi^2=5{\cdot}69$	—	$\chi^2=8{\cdot}91$
$P<0{\cdot}001$	$P>0{\cdot}20$	$P<0{\cdot}02$	$P=0{\cdot}14$	$P<0{\cdot}01$
$\phi=\cdot424$	$\phi=\cdot163$	$\phi=\cdot323$	—	$\phi=\cdot243$
N=264	N=45	N=45	N=87	N=87

In a simple comparison between grammar and non-grammar school the expected relation was strong. (Table 6). Thirty-five out of the fifty-three parents who had attended grammar school have successful children. Controlling on social class for both husband and wife produces an interesting suggestion which was considered earlier, but for which there was insufficient evidence. Then it was found that the previous *occupation* of the wife did not seem to have the relationship with working class success which is usually expected. Here a relationship does exist between the mother's *schooling* and success. It is plain that grammar school mothers have a lot to do with working-class success since half of the grammar school mothers in the working-class had successful children. This is a success rate almost five times as great as that found in the working-class as a whole. However, the small numbers involved undermine the validity of the ratio.

Within the middle-class only three out of seventeen grammar school women have unsuccessful children. There are too few grammar school husbands in the working-class for a significant relationship with success to show and in the middle-class they did not do as well as might have been expected. One explanation for this latter fact may be found in the idea that middle class life allows the 'potentiality' of the child to be exhibited. Assuming that the tests are accurate samplings of middle-class life, the performance of middle-class children will represent individual differences of a genetic nature. There is another explanation however—that differences in family ideology will be important and this will be taken up in a later paper.

Fraser found correlation coefficients of 0·42 and 0·49 between parental background and child performance on I.Q. and criterion. The difference between these correlations was then interpreted to mean that, given similar potential intelligence, children of the higher educated parents would perform better on the I.Q. and attainment tests. This is corroborated by the present study as far as the working-class is concerned. In the middle-class, parental educational level does not seem to have the same importance.

Parents' Attitudes Towards Education. In an effort to quantify parental attitudes towards education an index was constructed which was very strongly related to success in the working-class. Overall, it was also very much an aspect of ' middle-classness " (Table 7).

A mark was given on each of the following circumstances :
If grammar or public school was preferred for child.
If *both* parents expect child to leave school at 18 years.
If *both* parents chose a satisfactory attitude towards homework.
If *both* parents chose the educational ladder as the best way to get on in life.
If *both* parents chose 18 years as minimum schooling necessary.
If *either* parent attended at least one year of academic night school.
If head teacher rates ' attitude towards the educational experience of the child '
A or B on a five-point scale.

TABLE 7

THE RELATIONSHIP BETWEEN PARENTAL ATTITUDES TOWARDS EDUCATION AND CHILD'S SUCCESS ON THE 11+ OR MEMBEPSHIP OF THE MIDDLE CLASS.

Index of Attitudes towards Education Associated with :			
11+ Success			Membership of Middle Class
Within Total Sample	Within Middle Class	Within Working Class	Within Total
$\chi^2=36\cdot32$	—	$\chi^2=18\cdot20$	$\chi^2=34\cdot42$
$P<0\cdot001$	$P=0\cdot17$	$P<0\cdot001$	$P<0\cdot001$
$\phi=\cdot524$	—	$\phi=\cdot408$	$\phi=\cdot511$
N=132	N=45	N=87	N=132

Clearly the members of the two categories are very different in their attitudes toward education, but the picture alters when they are compared in the different social classes ; for, whilst the relationship is very strong in the working class ($\phi=\cdot408$) it appears to be much less so in the middle-class. Scores above 4 are so much a middle-class phenomenon ($\phi=\cdot511$) that it is very hard for a tendency to show itself in the middle-class.

Dealing with the working-class first, we can say that, as all other studies have found, parents whose views of education are sympathetic and who see education as an important ladder for mobility, tend to have children who are successful in school. This does not prove that the attitudes of these parents ' produce ' such children any more than it proves that such children encourage parents to think in that sort of way. But, bearing in mind the fact that, at the time of interviewing, the parents did not know for certain whether or not their child would be allocated to grammar school, a genuine, mutually reinforcing relationship between parental attitudes towards education and the child's ' ability ' seems undeniable.

Parents were also asked how they would feel if their child were not awarded a grammar school place on the selection procedure. The forced-choice answers of those who actually wanted them to go to grammar school took up an interesting pattern. (Table 8). Here there is a clear difference within the middle-class :

TABLE 8

THE RELATIONSHIP BETWEEN PARENTS' ANTICIPATED RESPONSE TO ' FAILURE ' ON THE SELECTION PROCEDURE AND EITHER THE ACTUAL RESULT OR MEMBERSHIP OF THE MIDDLE CLASS. PARENTS WHO PREFERRED THEIR CHILD TO GO TO GRAMMAR SCHOOL.

' Would not Mind ' Associated with :			
Failure of Child on 11+			Membership of Working Class
Within the Total Sample	Within the Middle Class	Within the Working Class	Within the Total Sample
$\chi^2 = 15{\cdot}68$	$\chi^2 = 6{\cdot}40$	$\chi^2 = 2{\cdot}74$	$\chi^2 = 10{\cdot}04$
$P < 0{\cdot}001$	$P < 0{\cdot}02$	$P < 0{\cdot}10$	$P < 0{\cdot}01$
$\phi = {\cdot}277$	$\phi^2 = {\cdot}276$	$\phi = {\cdot}151$	$\phi = {\cdot}222$
$N = 204$	$N = 84$	$N = 120$	$N = 204$

A significantly greater proportion of unsuccessful-to-be parents were prepared to say that they would not mind if their child was not sent to the grammar school under the allocation procedure. This can hardly be taken at its face value.

In the first place, middle-class parents are apt to be accurately appraised of the chances of their children in the ' 11+ ' Consequently, they have a better opportunity and more time to adjust their feelings to suit the situation. In the second place, this is not the only avenue to a grammar school type of education—only the cheapest. Finally, within the middle-class there was a surprisingly large number of people—usually teachers, or people who have some sort of connection with the world of education—who believed in the ' age, aptitude and ability ' approach to educational guidance. Here the genuine answer really was that they did not mind, because it was best for the children.

The ' would not mind ' group within the working-class are interesting because the successful parents all responded in this way for what was otherwise a middle-class reason—' the educationalists know best '. On the other hand, the principle reason amongst the ' failed ' working-class group was the expected one of apathy and indifference to something which was outside ' our ' sphere anyway. After all this, it is perhaps not surprising to find that membership of the working-class associated with acceptance of possible failure ($\phi = {\cdot}22$). Remembering also, the parents of the working-class who actually chose the secondary modern school (fifty-six) and who were certain to be satisfied, it is plainly true to say that disatisfaction with the prospect of being allocated to a modern school is very much a middle-class phenomenon.

Within the classes it can be seen that the tendency to expect some form of disappointment, should the child fail, is much stronger in the middle-class. This is exactly what one would expect. To the middle-class, ' our ' school is the grammar school. It is taken for granted. To the working-class ' our ' school is still the secondary modern school. Even amongst the successful-to-be parents, the grammar school is something to aspire to, not to expect ; so that a factory supervisor has to qualify her choice of a minor public school with an embarrassed giggle : " that's only one in a million for poor class children." " It's a school that people from it have always been a little higher—more brainy."

IV.—CONCLUSION.

This paper has dealt with some evidence relating to the connection between certain aspects of the family environment and school performance. Each factor has been included because it portrays an aspect of the environment which is thought to be important in school performance. The intention has been to emphasise that we have some way to go yet in distinguishing between genuine causal factors and purely statistical coefficients which derive their predictive efficiency (such as it is) from their relationship with certain configurations of socio-cultural elements.

Analysis of these configurations is still at a crude level which employs occupational criteria in categorising family social environment. Because of this, it is important to understand what is being attempted when data are analysed according to social class criteria. Occupations are usually used for this because they introduce a simple criterion upon which may be based an assumption of similarity of work- and life-experiences. It is assumed that from this similarity will follow a tendency towards similar ways of life, which are expected to stimulate the development of similar cognitive skills and habits. A category called ' middle-class ' is put together on the assumption that the resulting group will tend to have similar values, life-orientations and ways of thinking which are interdependent, and collectively different from those characterising working-class groups. (Kohn, 1963). These different ' patterns of culture ' are responsible, to an unknown extent, for the variations in intellective skills and habits of children. (Triandis, 1964). How soon it will be before we can describe clearly the mechanisms by which the different patterns bring about the different abilities will depend as much upon improvements in our ability to measure variation in the cultural environment as upon improvements in our ability to measure differences in exhibited intellective skills.

V.—REFERENCES.

BARBER, B. (1961-62). Social-class differences in educational life-chances. *Teachers' College Record*, 63, 102-113.

DOUGLAS, J. W. B. (1964). *The Home and the School.* London : Macgibbon and Kee.

ELDER, G. H. (1965). Life opportunity and personality : some consequences of stratified secondary education in Britain. *Soc. of Educ.*, 38, 173-202.

EYSENCK, M. J., and EYSENCK, S. B. G. (1964). *Manual of the Eysenck Personality Inventory.* London : Univ. of London Press.

FLOUD, J. E., MARTIN, A. M., and HALSEY, A. H. (1956). *Social Class and Educational Opportunity.* London : Heinemann.

FRASER, E. (1959). *The Home Environment and the School.* London : Univ. of London Press.

HALSEY, A. H. (Ed.) (1961). *Ability and Educational Opportunity.* Paris : Organisation for Economic Co-operation and Development.

HERRIOTT, R. E. (1963). Some social determinants of educational aspiration. *Harvard Educ Rev.*, 33, 157-177.

KOHN, M. L. (1963). Social class and parent-child relationships : an interpretation. *Amer. J. of Sociology*, 69, 471-480.

NISBET, J. (1953). Family environment and intelligence. *Eugenics Rev.*, 45, 31-42.

ROGOFF, N. (1960). Public schools and equality of opportunity. *J. of Educ. Sociology*, 33, 252-259. 0

SWIFT, D F. (1965). Meritocratic and social class selection at age 11. *Educ. Res.*, VIII (2), 65-73.

—— (1966). Social class and achievement motivation. *Educ. Res.*, VIII (2), 83-95.

—— (1967). Social class, mobility-ideology and 11+ success. *Brit. J. Sociology* (forthcoming).

TRIANDIS, H. A. (1964). Cultural influences upon cognitive processes, in BERKOWITZ, L. (Ed.). *Advances in Experimental Social Psychology, Vol. I*, New York : Academic Press, 2-48.

(*Manuscript received, 14th October*, 1965.)

R 71.153

THE TEACHER'S ROLE: A COMPARATIVE STUDY OF THE CONCEPTIONS OF COLLEGE OF EDUCATION STUDENTS AND HEAD TEACHERS

By DOUGLAS S. FINLAYSON
(*University of Liverpool Institute of Education*)

AND LOUIS COHEN
(*Anstey College of Physical Education*)

SUMMARY. A role definition instrument was constructed consisting of twenty-two statements of expectations for teacher behaviour in four role sectors—organisation, general aims, motivation, and classroom management; 268 female student-teachers and 183 head teachers of primary and secondary schools completed the role definition schedules. Mean response scores and measures of consensus were computed. Students were categorised by year of training and professional destination. Comparisons of responses within the student sample and between students and head teachers were made. Few significant differences were found between teachers of younger children and teachers of older children. A consistent pattern of change towards less authoritarian classroom behaviour was found to reach its peak in the second year students' responses. More traditional views were found in the third year group. In the student-head teacher comparison widespread disparity between role expectations was found to exist in all four role sectors. These conflicting student and head teacher role expectations are discussed in terms of a college and a school frame of reference, and a dual-phased process of teacher training.

I.—RESEARCH PROCEDURE.

THE few empirical studies which have been done in this country of the role of the teacher (Musgrove 1961, Taylor 1962, Musgrove and Taylor 1965) have assessed the values which teachers might be expected to have in relation to educational aims. This study utilised expectations of their actual behaviour in school situations. The aim was to examine the changes which took place in student teachers' conceptions of the teacher's role as they progressed through their course of professional training, and to compare the expectations of the student teachers with those of head teachers.

A role is defined by Gross (1958) as a set of expectations applied to an incumbent of a position in a social system. In this study such expectations were measured by a role definition instrument (R.D.I.)* which contained a number of statements about teacher behaviour derived from the literature and students' essays. To reduce response set, desirable and undesirable statements were used. The statements related to four role sectors discussed by Fleming (1958) which were judged suitable for initial exploration:

(i) *Organization.* These items included expectations about the way children were grouped in the school, and about the clerical and supervisory duties of the teacher.

(ii) *General Aims.* These items included expectations for the teacher's work in such areas as teaching the three R's, inculcating attitudes and values, and helping children to acquire good manners and correct speech. Controversial curriculum content areas such as Sex Instruction and Religious Education, were also included in this sector.

---

* A copy of the R.D.I., with the items categorised into these four sectors, is given in the appendix.

(iii) *Motivation.* These items included expectations for teacher behaviour as the motivator of children's learning, e.g., by comparing the work of different children in the class. The extent of belief in the 'discipline' of uninteresting work was also assessed.

(iv) *Classroom Behaviour.* These items included expectations about the teacher's personal relationships with the children, his approach to individual children, the kind of social climate he sought to establish, and the mode of leadership that he favoured.

Respondents were required to indicate for each item which of five modes of response best expressed their own expectations about that particular item, viz., that the teacher absolutely must, preferably should, may or may not, preferably should not or absolutely must not behave in this way. These response categories were weighted with scores from 1 to 5.

The R.D.I. was submitted to the entire student population (268 females) of a college of education. It was also given to a sample of 183 head teachers, made up of almost equal numbers from primary and secondary schools. The primary head teachers were selected from school lists of five divisions of one county authority, and three county boroughs. The secondary head teachers were selected from the school lists of one county authority, from four county boroughs and two excepted districts. 61 per cent. of the primary, and 55 per cent. of the secondary head teachers returned completed schedules.

II.—STATEMENT OF RESULTS.

(i) *Differences between types of students.* Initially, the mean scores of students preparing to teach infant children were compared with those of students training to teach junior and secondary children. Of all the items on the R.D.I., the '*t*' test showed significant differences above the 5 per cent. level on only four. These were concerned with classroom management. The students training to teach older children showed more authoritarian attitudes. They stressed the need for children to acquire better manners and correct speech, favoured more punitive methods of control, and expected to maintain a more distant and formal relationship with their pupils than did the teachers of infant children.

(ii) *Differences between year-groups.* The responses to each of the items by the three year-groups of students were treated by analysis of variance. Table 1 shows the means of each of the three groups on the items grouped according to role sectors. F ratios have been calculated only where the mean sum of squares between groups is larger than the mean sum of squares, within groups, and only for those items which produced significant F ratios are the differences between the means shown. Where the subsequent application of the '*t*' test showed these not to be significant this is indicated.

Of the items in the organizational category, only in relation to the peripheral duties expected of teachers such as the collection of savings and the supervision of school meals is any significant change noted. There is an increasing trend to reject these expectations as teacher functions as the students progress through college. There is, too, a growing opposition to the notion that children should be grouped by academic attainment, but the differences do not reach the ·05 level of significance.

TABLE 1

ITEM MEANS, F RATIOS and DIFFERENCES BETWEEN STUDENT YEAR GROUPS.

Item No.	1st Year Mean ($\bar{X}_1$)	2nd Year Mean ($\bar{X}_2$)	3rd Year Mean ($\bar{X}_3$)	F Ratio	Difference between Means		
					$\bar{X}_1-\bar{X}_2$	$\bar{X}_2-\bar{X}_3$	$\bar{X}_1-\bar{X}_3$
1	3·206	3·413	3·651	5·683†	—·207	—·238	—·445
2	3·686	3·835	4·000	1·365			
3	1·618	1·450	1·512	<1			
4	1·324	1·525	1·407	2·911			
5	2·824	2·875	2·756	<1			
6	3·265	3·125	2·721	4·149*	·140 NS	·404	·544
7	1·490	1·438	1·628	1·495			
8	2·382	2·325	2·233	<1			
9	2·693	2·788	2·894	<1			
10	3·832	4·013	3·953	<1			
11	2·382	2·588	2·186	3·250*	—·206	·402	·196
12	3·725	4·063	3·826	1·632			
13	2·775	3·538	2·872	13·549†	—·763	·666	—·097 NS
14	2·970	3·487	2·907	4·945†	—·517	·580	·063 NS
15	1·765	2·544	1·942	11·426†	—·779	·602	—·177 NS
16	2·363	2·400	2·291	<1			
17	2·275	1·913	2·023	4·571*	·362	—·110 NS	·252
18	2·455	2·650	2·776	1·434			
19	1·647	1·575	1·477	1·544			
20	2·470	2·975	2·791	3·932*	—·505	·184 NS	—·321
21	2·931	3·713	3·209	9·258†	—·782	·504	—·278
22	3·343	2·963	3·198	4·088*	·380	—·235	·415 NS

*—Significant at the 5 per cent. level.
†—Significant at the 1 per cent. level.

Of the items dealing with the general aims of education, again only one item shows differences large enough to be significant. This is the item dealing with the relative importance of the basic subjects. As the students progress through college, they tend to regard these as increasingly important. Another item which just fails to reach the level of significance is that dealing with the acquisition of correct speech and manners, where the second-year student appear less mandatory in their expectations than either the first or third year students.

Neither of the items dealing with the motivation of children shows any significant changes throughout the course of training.

Of particular interest are the items dealing with classroom behaviour. In this section eight out of the twelve items showed differences large enough to be significant, half of them at the 1 per cent. level. Not only are there more items recording significant changes, but also these changes conform to a consistent pattern. A study of the group means suggests that, in the second year, the students hold beliefs that indicate less authoritarian views about the pupil-teacher relationship and give stronger support to practices promoting the mental and emotional health of their pupils, than either first or third year students. The items which significantly demonstrate this attitude of growing liberality which then gives way to more traditional views are those in which the second-year students see themselves as: (*a*) less of a judge of right and wrong for the children ; (*b*) less punitive towards the aggressive child ; (*c*) more consistent in their classroom behaviour ; (*d*) more approachable ; (*e*) more prepared to allow children to learn from their own experience ; (*f*) less concerned with immediate obedience ; (*g*) more prepared to create warm, harmonious relationships in the classroom.

Item 12 shows the same trend towards a peak of permissiveness in the second year, but the differences do not reach the level of significance. Amongst the remaining three items in this category, item 16 shows scarcely any differences at all between years, and both items 18 and 19, although not sufficient to be significant, show a growing acceptance throughout the three years of a supportive role for the teacher, in which she sees herself more as a confidante and more prepared to show affection towards children in her class. Both of these items, however, deal with relationships with individual children rather than with the class *en masse*.

(iii) *Students and head teachers*. A comparison of the mean responses of students, the three year-groups being taken together, with those of the head teacher sample was now carried out and the results are shown in Table 2.

TABLE 2

ITEM MEANS, VARIANCES AND DIFFERENCES BETWEEN STUDENT AND HEAD TEACHER GROUPS.

Item No.	Students		Head Teachers		Difference Between Means	Level of Significance
	$\bar{X}$	σ^2	$\bar{X}$	σ^2		
1	3·410	·839	2·697	1·375	·713	·001
2	3·831	·957	3·038	1·070	·793	·001
3	1·534	·682	1·764	·793	—·230	·01
4	1·411	·317	1·094	·114	·317	·001
5	2·817	·993	2·175	·590	·642	·001
6	3·049	·1777	2·551	1·672	·498	·001
7	1·519	·585	1·458	·399	·061	NS
8	2·317	·784	3·535	·939	1·218	·001
9	2·786	1·580	2·112	·769	·674	·001
10	3·925	1·197	3·406	·996	·519	·001
11	2·381	1·042	1·425	·518	·956	·001
12	3·858	1·603	3·324	1·958	·534	·001
13	3·034	1·495	2·077	1·028	·957	·001
14	3·102	1·722	2·980	1·583	·112	NS
15	2·052	1·346	1·863	·763	·189	·05
16	2·351	·675	2·953	·760	—·602	·001
17	2·086	·713	2·762	1·029	—·676	·001
18	2·617	1·703	2·639	1·702	—·022	NS
19	1·571	·439	2·287	·903	—·716	·001
20	2·726	1·537	1·894	·998	·832	·001
21	3·254	1·577	3·068	1·358	·186	NS
22	3·183	·817	3·024	·594	·159	·05

As the mean responses of the primary and secondary head teachers were very similar, they have been combined for purposes of comparison with the students. The 't' test showed that 82 per cent. of the differences between head teachers and students were significant.

Each of the three organizational items showed a significant difference, the head teachers being much more in favour of teachers performing peripheral duties, more in favour of grouping children by academic attainment and less in favour of the teacher having access to children's records.

In the general aims category, the only item on which the head teachers and students showed a similar level of agreement was that concerned with helping children to acquire attitudes and values not fostered in their homes. On all the other items in this category, highly significant differences are shown. The head teachers are, on the one hand, more in favour of teachers giving religious education and, on the other, less in favour of their giving sex instruction than the students. In relation to manners, dress and speech, the head teachers hold more mandatory views than the students and also give greater importance to the teaching of the three R's.

Both the motivational items show highly significant differences, the head teachers believing more in the alleged discipline of uninteresting work and being less opposed to the comparison of one child's work with another.

In the behavioural item comparisons, those which did not reach the ·05 level of significance related to expectations about a consistent classroom discipline, the showing of affection to children and the giving of praise. On all the other items, significant differences are demonstrated. Head teachers are much more strongly in favour of teachers interpreting right and wrong for the child, using punishment, being stricter in their discipline, and insisting on immediate conformity from children than are the students. On the other hand, they are more inclined to reject activites in which children formulate their own rules of behaviour, opportunities for children to learn from their own experiences and for them to discuss their personal difficulties with teachers. In all of these items, the head teachers' greater concern for good order and discipline, and outward conformity is clearly manifested.

The disparity which has been demonstrated between the role conceptions for the teacher held by the students and the head teachers is extremely widespread. It exists in all four role sectors, and to some extent, is a measure of possible conflict for these students as they move into the schools as teachers.

In an endeavour to clarify those areas in which conflict is most likely between students and head teachers, the items were listed in rank order of mean score for each group separately and grouped according to whether the mean responses placed them in the 'agree' category, the 'uncertain' category or the 'disagree' category. By comparing the groupings of head teachers and students separately, it is possible to pick out those items which the two groups place in different categories. These items are listed in Table 3 below.

TABLE 3

GROUPING OF ITEMS ACCORDING TO RESPONSES OF HEAD TEACHERS AND STUDENTS.

	Agree	Uncertain	Disagree
Head Teachers	9, 13, 5, 20	16, 17, 12, 2	8
Students	16, 17, 8	9, 13, 5, 20	12, 2, 10

As will be seen, these items represent all the four role sectors but two general types of items can be distinguished. There are the items having motivational implications and those which involve considerations of authority, discipline and value judgment. In the motivational group of items, the head teachers are uncertain about placing slow learners with slow learners and about comparing one child's work with that of another child, whereas the students disagree with these statements. In the set of items concerned with the use of authority the students are uncertain about the discipline of uninteresting work, the punishment of aggressive children, the teaching of religious education, and seeking to get children to unquestinoningly obey orders whereas the head teachers agree that these things should be done by a teacher. The head teachers are uncertain about caning, about the formation of class councils and allowing children to act upon what teachers consider to be wrong decisions, whereas the students reject caning, and are positively orientated towards class councils and allowing children the opportunity to learn from their own mistakes. The item involving sex instruction is particularly conflicting, the students feel that the teacher should do this; the heads do not.

An interesting point can be made about the relative importance of educational aims held by teachers. The only item which students regarded as mandatory related to helping children to acquire correct manners and speech. The head teachers also placed this item first in their responses. Along with it in the mandatory category was their view that teachers should help children acquire attitudes and values not fostered in their own homes and this item came second in the student's rank order. Thus, both groups regard the function of social training very highly indeed. In fact, they regard it as more important than giving first priority to teaching the three R's. This finding is in contrast to that of Musgrove and Taylor (1965) and the explanation may well lie in the differences in the samples, or, more likely, in the way in which the general aims were presented for evaluation. In Musgrove and Taylor's study the teachers were required to rank the general aims in order of importance, whereas in this study, specific behavioural descriptions tend to be used. In this way, our results probably convey the popular notion that, until a class accepts a certain measure of social conformity, no formal instruction is possible.

The variance of the responses from any group is an indication of the consensus amongst the individuals. The two items on which the two groups were in greatest agreement amongst themselves were those just discussed, relating to social training. To study consensus further, the items were arranged in rank order, according to variance, each group being treated separately. A median split in the rank order of both groups was made to distinguish low from high variance items. The items which had low variance in both groups were those statements which contained general aims or dealt with organizational matters. Thus, there was greatest consensus amongst the groups about the extent to which teachers should encourage children to acquire good manners, correct speech, and appropriate attitudes and values, teach religious knowledge, group friends, allow the formation of a class council, have access to record cards and give sex instruction. There was also high consensus about pupil-teacher relationships which would encourage child confidences. The items for which both groups had high variances were generally those dealing with face-to-face classroom situations, such as comparing the work of children, obedience, punishment, the use of praise, consistent classroom behaviour, the showing of affection, and the priority to be given to the teaching of the three R's, which presumably, would depend on a number of factors in any particular class.

III.—DISCUSSION AND CONCLUSIONS.

When a student begins a three-year course of teacher-training, it might be hoped that this would be the start of a growth process in which the student becomes increasingly aware of all the aspects of the teacher's role and makes these expectations part of his own conception of teacher behaviour. The data discussed in this paper suggest that no matter how gratifying such a view might be to those engaged in teacher training, it is not tenable on several grounds:

(*a*) Firstly, when one looks at the table of differences between the first and third-year groups of students what stands out is not the number of differences but the number of similarities in mean responses of the first and third years. In 72 per cent. of the items there are no significant differences. In the six items in which there are differences, one is an increasing resistance to peripheral duties which, presumably, is not the result of college influence, and another is an increasing tendency to give the three R's a higher priority. There are, too, some surprising absences of changes. For example, there is no relinquishing of the faith in the discipline of performing uninteresting tasks, no increasing awareness of the need to maintain a consistent pattern of classroom behaviour, and no reduction in the comparison of work belonging to children. One would have expected that students would have become increasingly aware of motivational considerations and of the necessity of classroom consistency during their training and that this awareness would have produced some changes of expectancy. Thus, the general impression that one takes away from this first and third year comparison is the relative absence of difference between them.

(*b*) A second reason for maintaining that teachers in training do not acquire growing insight into the teacher's role can be found in changes which take place between the first and second years of the course and the second and third years. As was pointed out, the role sector in which there is the greatest change is that dealing with the classroom expectations for the teacher. The nature of the changes in the student's conceptions of the teacher's role suggests that there is a high-peak of understanding and insight which occurs during the second year. Less authoritarian attitudes and pupil-teacher relationships which pay more regard for practices and beliefs concerned with the mental health and emotional well-being of the children characterise this period but, by and large, these trends give way in the third year to beliefs and expectancies similar to those which students had at the beginning of their course of training, and which they gained, no doubt, from being pupils at school. The second year of training is, in fact, the time during which the student is furthest away from the school situation as pupil or teacher, and the findings of this study suggest that this time coincides with the students' maximal liberality.

(*c*) The third reason for rejecting the notion of steady growth in understanding arises from a consideration of the head teachers' conceptions of the teacher's role. When their conceptions are compared with those of the students, what is obvious is the very high percentage of significant differences, most of them at a high level. The direction of these differences show that the head teachers' conception of the teacher's role is generally in a contrary direction to the changes noted in the first two years of the students' training. In fact, the head teachers' conceptions go beyond the mean responses of the student-teachers' conceptions towards a relatively organizationally-orientated, child-dominating, and conformity-desiring point of view. A hypothesis is suggested that this predominantly organizational focus of the head teachers' expectations for his staff may well stem from his occupancy of the position of ultimate responsibility for these matters, but this hypothesis would require to be tested. Mention has already been made of the possible conflict areas which face the student-teacher as he enters school as a teacher. As the disparity between the students' and head teachers' conceptions of the teacher's role is maximal during the second year of training, it could be that the regressions in expectations which take place during the third year of training are evidence of the students' growing awareness of these disparities and that they are seeking to narrow the gap in the conceptions of what is thought desirable teacher behaviour by the college and what they see will be expected of them in schools.

Further support to this argument is given by the fact that the role sector in which these regressions of conception occur is that of classroom behaviour. This is the role sector in which expectations are most obviously reflected in behaviour and hence most open to observation and evaluation by others. The fact that the extent of change in any of the other role sectors is very slight indeed, perhaps suggests that in these other sectors beliefs are less subject to pressure from the organizational demands made by the head teacher. A teacher can, for example hold beliefs about school organization and curricular content without ever disclosing them for they will not be reflected in his behaviour as these areas of decision-making tend to be the responsibility of the head teacher, not the individual teacher. Regarding expectations about motivating children, these may be reflected in a teacher's behaviour but their implementation need not necessarily affect the good order or organizational efficiency of the classroom: their effects would rather be detected in the feelings of individual children, and hence not open to the judgment of the head teacher.

Tentatively, a second hypothesis can be formulated which might be tested further, viz., that tutors in colleges of education may have one frame of reference.* from which to view the teacher's position while head teachers in schools have another.

(i) *College of Education frame of reference.* The college staff and students are detached from the everyday workings of school. Their positions do not have associated with them any responsibilities for running a school and for dealing with the variety of problems which arise in it. Their concern is perhaps more with learning about children and teaching rather than working with children in face-to-face relationships. The student thus, when asked to respond to the R.D.I., probably does so in terms of this contextual system.

(ii) *School frame of reference.* Teachers in schools are involved in dealing with children often under very adverse conditions, in which large classes, overcrowded conditions and lack of appropriate material are often featured. The head teacher has responsibility for the educational results of these teachers and there is little need to reiterate the complaints of too few teachers, untrained teachers and the frequent changes of staff which occur in many schools. It is in this context that the head teachers probably make their judgments of expected teacher behaviour. It is hardly surprising, therefore, that they tend to give high place to those aspects of teacher behaviour which emphasise the maintenance of good order and discipline rather than the individual growth and mental well-being of individual children.

Some evidence which supports this hypothesis of the two frames of reference has come from a comparison of the responses of a national sample of college of education tutors with the responses of the student group and those of the head teachers. Broadly speaking, the differences between the groups indicate a greater measure of agreement between tutors and students on the one hand, than between tutors and head teachers on the other.

---

* Krech, Crutchfield and Ballachey (1962) define a frame of reference as the contextual system with respect to which an object is judged. The object in this study is the teacher's position.

This hypothesis of the two frames of reference appears closely related to Gross' (1965) idea that taking up a professional appointment in schools should be considered as part of a two-phased process of socialisation. The first or preparatory phase is the period of formal training when the skills, knowledge, values and attitudes prescribed for entry into the post are taught. Presumably, during this time, students will have internalised an ideal conception of their role as a teacher. The second phase of training, that of organizational reality, begins when the student confronts the complex realities of organizational life and finds that theory and practice are rather different things.

While it may be necessary to think of teacher training in such terms, it has been pointed out that differences beyond a certain point in the role conceptions of these two aspects of the training process may imply conflict for the new teacher. Further implications stemming from these differences could be drawn relating to school-college relationships, to the role of inspectors and advisors, and to the rate and direction of change within the educational system.

In addition to these social implications, it is important to bear in mind the large variances reported for many of the items, particularly those relating to face-to-face classroom situations. This indicates the absence of any real consensus of opinion within groups about these expectations of teacher behaviour. Further work is obviously necessary to investigate the personality and other factors contributing to these differences. Another area which would require further investigation is the relationship between role conception and actual teaching behaviour in the classroom.

These comments indicate the limited nature of the present study of which various criticisms might be made. A longitudinal study would be preferable to a cross-sectional one. It would be preferable to draw a sample from a number of colleges and to include males as well as females. The response rate of the head teachers might have been improved. What is maintained, however, is that the results are of sufficient interest, and the implications which might be drawn of sufficient importance to justify a much bigger and more broadly based inquiry which seeks to integrate personality assessment, role conception and classroom behaviour.

IV.—REFERENCES.

Burnhah, P. S. (1964). *The Role of the Deputy Head in Secondary Schools.* M.Ed. dissertation, University of Leicester.

Butcher, H. J. (1959). *The Opinions of Teachers and Student Teachers about Education.* Ph.D. dissertation, University of Manchester.

Cannon, C. (1964). Some variations in the teacher's role. *Education for Teaching,* Vol. No. 64.

Cohen, L. (1965). *An Exploratory Study of the Teacher's Role as Perceived by Head Teachers, Tutors, and Students in a Training College.* M.Ed. dissertation, Liverpool University.

Fleming, C. M. (1958). *Teaching—A Psychological Analysis.* Methuen.

Gross, N , Mason, W. S., McEachern, W. B. (1958). *Explorations in Role Analysis : Studies in the School Superintendency Role.* New York : Wiley.

Gross, N., and Herriot, R. E. (1965). *Staff Leadership in Public Schools.* New York : Wiley.

Hollins, T. H. B. (1955). *Teachers' Attitudes to Children's Behaviour.* M.Ed. dissertation University of Manchester.

Kissack, M (1956). *The Attitude of Training College Students towards Corporal Punishment.* M.Ed. dissertation, University of Manchester.

Krech, D., Crutchfield, R. S., and Ballachey, E. L. (1962). *Individual in Society.* McGraw-Hill Book Co., Inc.

Musgrove, F. (1961). Parents' expectations of the junior school. *Soc. Rev.*, 9, 167-180.

Musgrove, F., and Taylor, P. H. (1965). Teachers' and parents' conception of the teachers' role. *Brit. J. Educ Psychol.*, XXXV, 171-178.

Sharma, C. L. (1963). *A Comparative Study of the Pricesses of Making and Taking Decisions within Schools in the U.K. and U.S.A.* Ph.D. dissertation, University of London.

Taylor, P. H. (1962). Children's evaluations of the characteristics of the good teacher. *Brit. J. Educ. Psychol.*, XXXII, 258-266.

Webb, J. (1962). The sociology of a school. *Brit. J. Soc.*, 13, 264-272.

Wilson, B. R. (1962). The teacher's role—a sociological analysis. *Brit. J. Soc.*, 13, 15-32.

(*Manuscript received 8th November, 1965*)

V.—Appendix.

The Role Definition Instrument.[1]

1.—Supervise school meals, collect savings, etc.

2.—Put ' slow ' learners with ' slow ' learners in all academic work.

3.—Have access to the children's personal record folders and to all information contained there.

……

4.—Help children acquire good manners, correct speech.

5.—Teach Religious Knowledge.

6.—Teach the three R's as his primary responsibility as a teacher.

7.—Help children acquire attitudes and values not fostered in them in their own homes.

8.—Give sex instructions to pupils.

……

9.—Alternate interesting work with less interesting so that the pupils will appreciate the former yet gain by the discipline of the latter.

10.—Use the comparison of one child's work with that of another as a method of motivation.

……

11.—Interpret ' right ' and ' wrong ' for children.

12.—Cane those children who need it.

13.—Punish the aggressive child for his attacks on other children.

14.—Never allow children to know how he will react to classroom situations.

15.—As a newly-qualified teacher, start as a strict disciplinarian and gradually become ' approachable ' as his class respects his authority.

16.—Encourage children to form a class council to make rules for their own classroom behaviour.

17.—Allow children at times to act upon what he thinks are wrong decisions on their part.

18.—Guard against showing affection to children in his class.

19.—Allow children to confide in him with personal problems that they may not wish to discuss with their parents.

20.—Teach children to obey orders at once and without question.

21.—Give praise sparingly lest it lose its effectiveness.

22.—Group friends together for work in Mathematics and English.

……

---

[1] The items above have been grouped according to role sectors. At the time of presentation, their order was randomised.

THE CHILD WHO DISLIKES GOING TO SCHOOL

By *SHEILA MITCHELL and MICHAEL SHEPHERD
(Institute of Psychiatry, London University)

SUMMARY. During the course of an enquiry into the behaviour of a random sample of over 6,000 children aged from 5–15 years, attending local authority schools in Buckinghamshire, it was possible to obtain information about each child's reaction to school. This indicated that dislike of school was associated with poor attainment at school and signs of anxiety at home. Boys who disliked school were also found to be significantly more likely to have a problem of behaviour reported by their teachers and were more unco-operative in class but this was not true of the girls. Reaction to school affected attendance among children of secondary school age, but not among those under 11 years. The implications of this are discussed in relationship to future research.

I.—INTRODUCTION.

THOUGH children aged from 5 to 15 years compulsorily spend much of their waking time in school, little information is available about their reaction to such attendance and the way in which it affects other aspects of their life at school and at home. Studies which have been carried out with the child as informant (Fitt, 1956; Wisenthal, 1965) have shown that boys were less favourably disposed to school than girls; that attitudes became less favourable with increasing age; and that attitudes were positively related to attainment.

In the present enquiry, information about reaction to school attendance was obtained not from the child but from the parents as one item among many included in a questionnaire concerned with the distribution of behavioural and emotional problems among a random sample of school age children.

II.—METHODS.

(1) *Coverage.* In June, 1961, questionnaires were sent to the parents and teachers of a one in ten random sample of children aged between 5 and 15 years who were attending local authority schools in Buckinghamshire. The main purposes of this enquiry were (i) to obtain information about the age and sex distribution of various types of behavioural and emotional problems in a random sample of children; and (ii) to relate behaviour to other aspects of the child's environment, both at home and at school. Completed questionnaires were received back from the parents of 93 per cent. of the children in the sample (approximately 6,300) and the teachers of 97 per cent. (approximately 6,600): completed forms from both teacher and parent were available for approximately 6,100 children (90 per cent.).

(2) *Assessment of Attitude.* In the questionnaire which was sent to parents, they were asked to assess the child's reaction to school by underlining the most appropriate sentence in the following triple-choice item:

Likes school very much.
Likes school about as much as most children.
Dislikes going to school.

It is apparent that this method of enquiry relates essentially to the parental percept of the child's outlook. The sentence underlined will depend on the parent's interpretation of the child's feelings from his behaviour and remarks:

---

* Now at the Department of Epidemiology and Preventive Medicine, The University of Glasgow.

for instance, whether " I don't want to go to rotten old school " denotes real disinclination or distress or simply a minor declaration of independence with little emotional content. The interpretation made will depend to a great extent on the parent's expectations about what type of behaviour would constitute the usual reaction of a child to his school and teachers. Some parents, particularly those who disliked school themselves, might consider a considerable amount of antipathy to educational procedures to be ' normal ' and underline the middle category in respect of a child whose reactions would cause a mother with more favourable memories to underline " Dislikes going to school." Parents who recorded the unfavourable category (5 per cent.), therefore, may be taken as comprising those whose children's reactions were strong enough to be considered atypical by their own family standards. Similarly, children who were said to like school very much (30 per cent. of the boys and 40 per cent. of the girls) would be likely to include those who showed active enthusiasm rather than mere acceptance. The bulk of the children (63 per cent. of the boys ; 55 per cent. of the girls), as might be expected, fell into the intermediate category.

III.—FINDINGS.

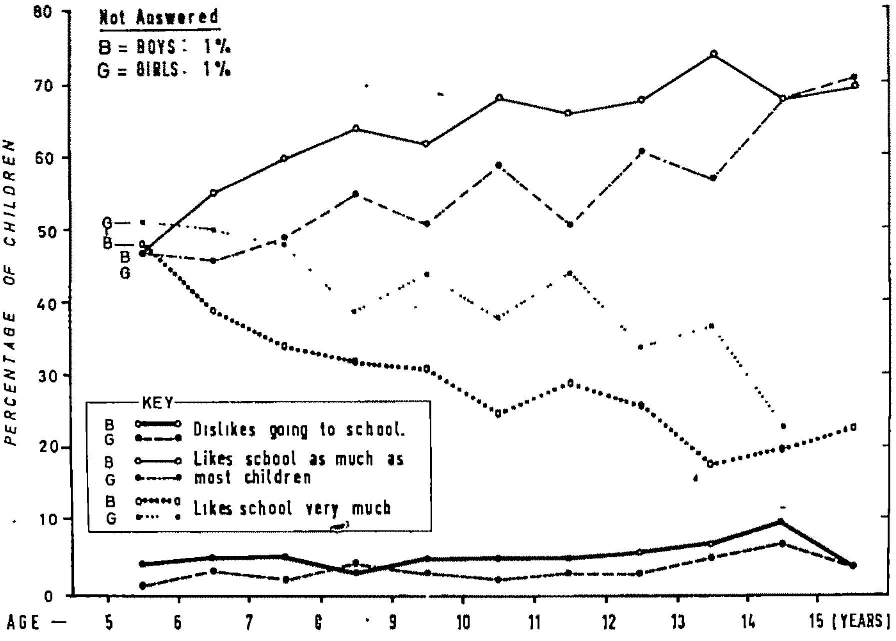


FIGURE 1

PARENT'S ASSESSMENT OF THE CHILD'S REACTION TO SCHOOL BY AGE AND SEX.

(1) *Age and Sex Distribution.* Figure 1 shows that a favourable reaction to school was more characteristic of girls than boys at all ages except the 15-year-olds. Conversely, more boys than girls were said to dislike going to school at all ages except 8 and 15 years. In each sex, there was a marked

tendency for enthusiasm for school to decline as age increased, though among the 11-year-olds there was a slight upward trend which may indicate a mild resurgence of enthusiasm accompanying the child's transfer to secondary school. Age, however, had little effect on the proportion of children who disliked going to school: this remained relatively constant at 5 per cent. of boys and 3 per cent. of girls up to the age of 12 years before rising among the 13 and 14-year-olds.

(2) *Attainment.* Within the school situation, the child who disliked going to school was markedly less successful than the child who liked school very much. Owing to the large number of schools involved in the survey it was impossible to obtain objective overall comparison of attainment. A rough measure of success in the child's own school environment was obtained, however, by asking teachers to rate the child's attainment relative to others of the same age, attending the same type of school. This showed a clear relationship between the child's attitude to school and the standard of attainment achieved (see Table 1).

TABLE 1

REACTION TO SCHOOL BY TEACHERS' RATING OF GENERAL ATTAINMENT LEVEL.

Level of attainment relative to others of the same age attending the same type of school	BOYS					
	Likes school very much		Likes school as much as most children		Dislikes going to school	
	No.	%	No.	%	No.	%
Above average	274	30	284	15	10	6
Average	421	47	904	48	63	40
Below average	205	23	697	37	86	54
Total	900	100%	1,885	100%	159	100%
Significance of association by χ^2 test	$p<0{\cdot}001$					

Level of attainment relative to others of the same age attending the same type of school	GIRLS					
	Likes school very much		Likes school as much as most children		Dislikes going to school	
	No.	%	No.	%	No.	%
Above average	375	32	271	17	10	10
Average	605	52	881	55	49	48
Below average	193	16	443	28	44	43
Total	1,173	100%	1,595	100%	103	100%
Significance of association by χ^2 test	$p<0{\cdot}001$					

Thus, among children who disliked going to school, the proportion rated below average in attainment was more than twice as great as that found among those who liked school very much: conversely the group who liked school very much contained a considerably higher proportion of children whose attainment was rated above-average than did either of the other two groups.

(3) *Attendance.* It might be expected that children described as disliking school would show a higher absence rate than those whose attitudes were more favourable. However, this did not seem to be the case in the 5–10 years age group. Thus, 20 per cent. of the boys in this age range who disliked going to school had, nevertheless, been credited with full attendance during the previous complete term: among the boys who liked school very much the proportion with full attendance was practically the same (22 per cent.). Similarly, 40 per cent. of the boys, in each additudinal group, were shown to have lost more than ten half-days of schooling. Among the girls aged 5–10 years, those who disliked going to school were less likely to have had full attendance during the previous term (7 per cent.) than were those with more favourable reactions (19 per cent. in each case). Overall, however, on the same type of distribution as shown for the older age group in Table 2, there was no significant* association between attendance and reaction to school among these younger girls.

TABLE 2

REACTION TO SCHOOL BY NUMBER OF SCHOOL SESSIONS LOST THROUGH ABSENCE IN THE PREVIOUS TERM, AMONG CHILDREN IN THE 11–15 YEARS AGE GROUP.

Parental Assessment of child's reaction to school	Number of half-day sessions absent in previous full-term							
	BOYS				GIRLS			
	0	1–4	5–10	11+	0	1–4	5–10	11+
Likes school very much	97 (35%)	70 (25%)	50 (18%)	58 (21%)	135 (32%)	105 (25%)	84 (20%)	101 (24%)
Likes school as much as most	211 (26%)	210 (25%)	117 (21%)	226 (27%)	180 (24%)	160 (22%)	162 (22%)	238 (32%)
Dislikes going to school	10 (12%)	20 (24%)	18 (22%)	34 (41%)	4 (6%)	13 (21%)	13 (21%)	33 (52%)

For each sex separately, by χ^2—test, $p<0{\cdot}001$.

Among the older children, on the other hand, this was not the case. If duration of absence is divided roughly into quartiles, as in Table 2, it becomes apparent that there was a significant relationship between the child's reaction to school and the number of sessions during which the child was absent. Thus, the proportion of children who had lost ten or more sessions during the term was twice as great among those who disliked going to school as among those who liked it very much: the proportion with full attendance decreased as the child's feeling about school became more negative.

(4) *Behaviour in School.* As can be seen from Table 3, there was a highly significant ($p<0{\cdot}001$) relationship among the boys between the number of behavioural items underlined by teachers and the child's feelings about school. Thus, in the older age group (aged 11–15 years), about one in two of the boys

---

* The chi-squared test was used throughout to assess significance: Yates' correction was applied in all 2×2 tables.

TABLE 3

RELATIONSHIP BETWEEN THE PARENTS' ASSESSMENT OF THE CHILD'S REACTION TO SCHOOL AND THE NUMBER OF BEHAVIOURAL ITEMS UNDERLINED BY THE TEACHERS.

Parental assessment of child's reaction to school	Percentage of children with				Total Number of children	Average No. of items underlined
	No items underlined	1 item underlined	2 items underlined	3 or more underlined		
A. Boys aged 5-10 yrs.						
Likes very much ...	53%	24%	13%	10%	648 (100%)	0·89
Likes as much as most	49%	27%	12%	12%	1,114 (100%)	0·99
Dislikes going	33%	27%	22%	18%	82 (100%)	1·44
Not recorded	35%	50%	8%	8%	26 (100%)	0·92
B. Boys aged 11-15 yrs.						
Likes very much ...	73%	19%	4%	4%	277 (100%)	0·42
Likes as much as most	67%	19%	8%	6%	837 (100%)	0·60
Dislikes going	45%	35%	10%	10%	82 (100%)	1·01
Not recorded........	40%	10%	30%	20%	10 (100%)	1·90
C. Girls aged 5-10 years						
Likes very much ..	62%	22%	9%	6%	775 (100%)	0·63
Likes as much as most	53%	24%	13%	10%	893 (100%)	0·89
Dislikes going	44%	32%	8%	4%	50 (100%)	0·73
Not recorded	33%	42%	4%	21%	24 (100%)	1·54
D. Girls aged 11-15 yrs.						
Likes very much ..	72%	20%	5%	3%	432 (100%)	0·42
Likes as much as most	67%	22%	8%	3%	751 (100%)	0·50
Dislikes going	61%	25%	6%	8%	64 (100%)	0·70
Not recorded........	47%	24%	6%	24%	17 (100%)	1·41

who disliked going to school had at least one behaviour problem underlined by the teacher compared with one in three of those who were said to like school " as much as most children " and one in four of those who liked it very much. (For a complete list of behaviour items covered in the questionnaire, see Appendix.) Among the younger boys the main difference was between those who disliked going to school and the rest. The distribution did not show any marked difference between the other two groups.

Among the girls, there was no significant association between reaction to school and the number of behaviour items underlined, though in the older age group, the trend was still visible. Among the girls aged 5 to 11, those who were said to dislike going to school showed fewer behaviour problems there than did those who liked it.

Only one specific item of behaviour in school was found to show an overall association with the child's reported attitude to school. As might be expected, this concerned lack of interest in school work (see Table 4).

It is true that many children in the group who were said to dislike going to school, were not reported as lacking interest by their teachers. Nevertheless, the proportion of such children was six to seven times as great among those who disliked school as it was among those who liked school very much and more than twice as large as among those who liked it ' as much as most children.' In each sex, the degree of association by the χ^2-test was significant at the 0·1 per cent. level.

TABLE 4

PERCENTAGE OF CHILDREN IN EACH ATTITUDINAL GROUP WHO WERE SAID BY TEACHERS TO BE " NOT INTERESTED IN SCHOOL WORK."

Parent's Assessment of Child's Reaction to School	Total Number of Children		Percentage marked by Teachers as " Not interested in school work."	
	Boys	Girls	Boys	Girls
Likes school very much	918	1,198	3%	2%
Likes school as much as most children	1,931	1,626	8%	5%
Dislikes going to school ...	161	106	21%	12%

Among the boys, there was also a significant ($p < 0{\cdot}001$) relationship between reaction to school and unco-operative behaviour while there. Thus, teachers marked as ' unco-operative in class ' 9 per cent. of those who disliked going to school compared with 2 per cent. of those who liked it very much and 4 per cent. in the intermediate category. Among the younger girls there was no association at all ; the proportion marked as unco-operative remained stable at 2 per cent. in all attitudinal groups. Older girls were more likely to be unco-operative if they disliked school (5 per cent. compared with 1 per cent. in the favourable group and 2 per cent. in the intermediate) but the degree of association was not significant.

No other behavioural items were reported by teachers to a significantly greater extent among those who disliked school than among other children in any age or sex group.

(5) *Behaviour at home.* It might be expected that children who did not like going to school would evince more signs of anxiety and, possibly, a greater tendency to psychological ailments at home. This appeared to be true to some extent. Thus, they were described in the parents' questionnaires as " Often seems worried, worries about many things " in 14 per cent. of cases—three times the proportion found among the other children (4 per cent.) : this difference is, of course, highly significant ($p < 0{\cdot}001$).

The children who disliked school were also significantly ($p < 0{\cdot}001$) more likely to have suffered from frequent headaches than were other children. Thus, 7 per cent. in the former group suffered from headache at least once a week compared with 3 per cent. of the latter. In all age and sex groups except the 5 to 10-year-old boys, more of the children who disliked going to school were also said to complain of frequent pains in their stomachs than were the rest of the sample. Thus, 13 per cent. of the girls and 9 per cent. of the older boys were said to suffer from such pains at least once in two weeks compared with 3 per cent. of the children who were more favourably disposed to school. Among the younger boys, however, there was a different picture : none of the children who disliked school had pains as often as once in two weeks while 4 per cent. of the other children did so.

(6) *Family Background.* It was only possible to examine the objective aspects of social back-ground : socio-economic status of the family, family size and the child's ordinal position in the family. In none of these was there a consistent relationship to the child's attitude to school.

TABLE 5

RELATIONSHIP BETWEEN THE CHILD'S REACTION TO SCHOOL AND THE OCCUPATIONAL STATUS OF THE FATHER.

Father's Occupational Status (Hall-Jones' Scale)	Child's Reaction to School					
	Likes school very much		Likes school as much as most children		Dislikes going to school	
	Boys	Girls	Boys	Girls	Boys	Girls
1, 2, 3. Professional, Managerial, Executive, Higher Grade/Supervisory	214	311	371	317	21	20
4, 5 (*a*). Supervisory, lower grade and routine non-manual	140	175	277	239	16	16
Total non-manual (No.) ..	354	486	648	556	37	36
(%)	40%	42%	35%	35%	24%	35%
5 (*b*). Skilled manual	296	388	711	582	58	45
6, 7. Semi-skilled and routine manual	235	283	518	444	61	22
Total manual (No.)	531	671	1,229	1,026	119	67
(%)	60%	58%	65%	65%	76%	65%

Among the girls, the group who liked school very much contained a significantly ($p<0.001$) higher proportion of those whose fathers were in non-manual types of employment but there was no difference between the socio-economic status of the fathers of those who disliked school and those who liked it " as much as most children." Among the boys, on the other hand, those who disliked going to school were significantly ($p<0\cdot001$) more likely to have had parents in manual occupations than were the other boys : those who liked school very much were still the most likely to be from non-manual homes, but, in this case, there was little difference from the boys who fell into the intermediate group.

So far as family size was concerned, girls who were only children were significantly less likely ($p<0\cdot05$) to dislike going to school (1 per cent.) than were those with sibs (4 per cent.) ; among the boys, the difference was small and non-significant but still followed the same trend (4 per cent. of only children : 6 per cent. of those with sibs). Among the children with siblings, there was no variation in the proportion disliking school as the size of the family increased.

Ordinal position in the family was also found to bear little relationship to the child's reaction to school. Among the girls there was no variation at all : the proportion who disliked going to school remained constant at 4 per cent. among eldest, middle and youngest children. Among the boys, however, those who were the youngest child in their family were significantly ($p<0\cdot01$) more likely to show disliked of school (7 per cent.) than were other children with sibs (4 per cent.).

IV.—DISCUSSION.

Several interesting points arise from this examination of the school lives of children whose parents described them as disliking to go to school. First, there is the strong relationship between unfavourable reaction to school and poor attainment level, which is in agreement with previous studies (Fitt, 1956; Wisenthal, 1965). It is difficult on the evidence available to be categorical about whether the child lacked success because of his unfavourable attitude to school or whether he disliked going to school because he was unsuccessful in school pursuits. It is noticeable, however, that among the children who disliked going to school, the proportion rated as below average in attainment was considerably larger than the proportion described by their teachers as lacking interest in school. It, therefore, appears possible that in some children dislike of going to school may be connected with dislike of lack of success in the class-room situation. Obviously, however, this is not the whole answer since many children of below average attainment were said to like school very much.

A second point of interest emerges from the lack of association between dislike of going to school and absence from school in the younger age group. If the child's dislike of school was a personal matter, it might be expected to lead to attempts by the child to use any excuse, such as a minor indisposition, to stay at home: if it reflected an anti-educational bias in the family background, one would expect parental support for such attempts. Yet, below the age of 11 years, the children who disliked going to school had an attendance record very little different from that of those who liked school. This would seem to indicate that, during the primary school years, children may achieve regular attendance under firm parental pressure and regardless of their own inclinations. This suggestion is supported by descriptions gathered from mothers during interviews carried out in connection with another aspect of the survey (Mitchell, 1965). Here several mothers of children with good attendance records told of difficulties they had experienced in ensuring that reluctant children reached school. Sometimes a parent or older child had to act as an escort as far as the school gate; in more extreme cases, where the child produced physical symptoms, usually vomiting, headache or abdominal pain, the mother had to steel herself to insist on school attendance, nevertheless—often in the face of adverse comments from neighbours. Such parental pressure appears to weaken as the child gets older and less amenable and this may account for the increased association between adverse attitude and poor attendance in the secondary school age group.

It seems that the existence of a group of children who dislike going to school but who show no indications of this in their attendance records is of importance for future research. Existing clinical studies dealing with the psycho-social associations of an adverse reaction to school attendance tend to limit themselves to school phobia, or 'separation anxiety,' and persistent truancy (Mullen, 1950; Tyerman, 1958; Davidson, 1960; Hersov, 1960; Kahn and Nursten, 1962) both of which involve actual failure to attend school. Future research, however, might usefully be extended to include those children who continue to attend school despite their adverse reactions. The inclusion of such a group might, for example, provide more comprehensive information about the relationship between maternal personality and separation anxiety since currently such explanatory hypotheses are based entirely on children who fail to attend school.

V.—REFERENCES.

CHAZAN, M. (1962). School phobia. *Brit. J. Educ. Psychol.*, 32, 209-217.

DAVIDSON, S. (1960). School phobia as a manifestation of family disturbance. *J. Child Psychol. and Psychiat.*, 1, 270-287.

FITT, A. B. (1956). An experimental study of children's attitude to school in Auckland, N.Z. *Brit. J. Educ. Psychol.*, 26, 25-30.

HERSOV, L. A. (1960). (1) Persistent non-attendance at school; (2) Refusal to go to school. *J. Child Psychol. and Psychiat.*, 1, 130-136, 137-145.

KAHN, J. H., and NURSTEN, J. P. (1962). School refusal: a comprehensive view of school phobia and other failures of school attendance. *Amer. J. Orthopsychiat.*, 32, 707-718.

MITCHELL, SHEILA (1965). *A Study of the Mental Health of School Children in an English County*. Ph.D. Thesis: University of London.

MULLEN, F. A. (1950). Truancy and classroom disorder as symptoms of personality problems. *J. Educ. Psychol.*, 41, 97-105.

TYERMAN, M. J. (1958). A research into truancy. *Brit. J. Educ. Psychol.*, 28, 217-225.

WISENTHAL, M. (1965). Sex differences in attitude and attainment in junior schools. *Brit. J. Educ. Psychol.*, 35, 79-85.

(*Manuscript received 10th January*, 1966)

VI.—APPENDIX.

BEHAVIOUR PROBLEMS INCLUDED IN THE TEACHERS' QUESTIONNAIRE.

(*a*) Very restless, can't sit still for a moment.
(*b*) Cries more than most children.
(*c*) Has a stammer.
(*d*) Has other speech difficulty.
(*e*) Often tells lies.
(*f*) Has stolen things on one or more occasions.
(*g*) Is very easily frightened.
(*h*) Bites finger nails.
(*i*) Sucks thumb or finger.
(*j*) Very irritable, easily becomes cross or annoyed.
(*k*) Has had one or more temper tantrums at school during this year.
(*l*) Is unco-operative in class.
(*m*) Very shy, finds it difficult to mix with other children.
(*n*) Has wet or soiled self at school during the last year.
(*o*) Has noticeable twitch of face or body.
(*p*) Worries more than other children.
(*q*) Aggressive towards other children.
(*r*) Very quiet or withdrawn.
(*s*) Very moody—on top of the world one minute, down the next.
(*t*) Not interested in school work.
(*u*) Has on one or more occasions during this school year shown fear of school—that is, tears on arrival or refusal to come into the building.

THE EFFECTS OF THE 'ELEVEN-PLUS' RESULT ON THE SUBSEQUENT CAREERS OF PUPILS

By E. D. CAVE

(Strabane County Secondary School)

SUMMARY. The circumstances and prospects of candidates in the I.Q. range 113–116, in Belfast and Down, were investigated eight years after the selection examination at 11 plus. A comparison was made between the 'qualified' group (i.e., those selected for academic grammar school education) and the 'unqualified' group in four social classes. It was found that the influence of social class was such that, especially at the extremes of the social class range, 'qualifying' or 'not qualifying' had little significant effect. In the middle of the social class range the 'qualified' pupils tended to stay longer at school, to have higher scholastic attainments and to reach a higher occupational level.

I.—INTRODUCTION.

IN 1955, the Qualifying Examination consisted of two Verbal Reasoning Tests, and two Attainment Tests in English and one in Arithmetic. The I.Qs. derived from the Verbal Reasoning Tests were used only in the allocation of the borderline group on the attainment test, i.e., those whose attainment lay between 250–299 marks, the qualifying mark being 300. One important and special factor in the Northern Ireland system of secondary education has been the provision for 'unqualified' pupils to enter grammar school as 'fee-payers,' and 48 per cent. of the 'unqualified' pupils in this investigation decided to pay fees. Using the higher I.Q. of the two obtained from the qualifying examination results, those candidates whose I.Q. was in the range 113–116 were chosen from Belfast and County Down for investigation. The range was chosen to ensure groups, roughly equal in numbers, of 'qualified' and 'unqualified' pupils whose I.Qs. did not differ significantly. A questionnaire was used to obtain information about the present circumstances of the candidates. Some difficulty was experienced in obtaining correct addresses as there had been considerable movements of population during the period 1955–1963. In all, 57 per cent. of the questionnaires sent out were returned, and the checks on these revealed no bias which would appear to affect the findings. One-hundred-and-sixty-seven 'qualified' pupils and 136 'unqualified' pupils made returns. A 't' test revealed no significant difference in the mean I.Qs. of the two groups. The two groups were each divided into four social classes on the basis of father's occupation. It was decided to use the following four broad social classes:

Class A. Upper and Middle—Professions, managers and directors, university teachers.

Class B. Lower Middle—Teachers, bank officials, civil servants, company representatives.

Class C. Skilled manual (and small shopkeepers).

Class D. Lower working.

TABLE 1

The Numbers in each Social Class in the 'Qualified' and 'Unqualified' Groups

	Numbers in each class			
	A	B	C	D
Qualified	14	70	58	25
Unqualified	20	40	51	25

II.—Some findings of the investigation.

(i) *Present Occupation.* As present occupational status was one of the most important of the criteria examined the findings are reported in some detail.

TABLE 2

The Distribution, with Respect to Present Occupational Level.

	Class	Present Occupational Level			
		'a'	'b'	'c'	'd'
Qualified	A	6	7	1	0
	B	13	45	12	0
	C	5	20	28	5
	D	1	3	17	4
Unqualified	A	5	14	1	0
	B	3	12	25	0
	C	2	8	36	5
	D	0	2	16	7

Classification of 'Present Occupational Level'.
'a' University students.
'b' Training college students; higher business and technical workers; civil servants; bank clerks.
'c' Skilled occupations.
'd' Unskilled, unemployed.

Only 16 per cent. of the 'qualified' pupils in class D reach an occupational level 'a' or 'b.' For classes C and D combined, the figure is 35 per cent. It is interesting to note the present occupations of the 'qualified' pupils in class D: clerks (6), apprentice radio engineers (3), apprentice machinists (3), compositors (2), painters (2), university students (1), student teachers (1) civil servants (1), comptometers (1), apprentice electricians (1), dock labourers (1), milk servers (1), shop assistants (1), millworkers (1). In class A, 95 per cent. of the 'unqualified' pupils reach occupational level 'a' or 'b.' Their present occupations are: university students (5), student teachers (4), nurses (3), bank clerks, (2), shorthand-typists (1), business trainees (1), student chartered accountants (1), make-up artists (1), heating engineers (1), student speech therapists (1). It is noteworthy that all these subjects entered grammar school as fee-payers.

For analysis, it was decided to combine social classes A with B, and C with D, and occupational levels 'a' with 'b,' and 'c' with 'd.' With both 'qualified' and 'unqualified' pupils a significant relationships was found between social class and present occupational level ($p < \cdot 01$ in both cases). (Table 3).

TABLE 3

THE PRESENT OCCUPTIONS OF THE 'QUALIFIED' AND 'UNQUALIFIED' GROUPS WITH SOCIAL CLASSES A and B, C and D COMBINED AND OCCUPATIONAL LEVELS 'a' and 'b,' 'c' and 'd' COMBINED.

'Qualified' Group.		
Social Classes	Present Occupational Levels	
	'a'/'b'	'c'/'d'
A/B	71	13
C/D	29	54
'Unqualified' Group.		
Social Classes	Present Occupational Levels	
	'a'/'b'	'c'/'d'
A/B	34	26
C/D	12	64

Table 4 shows that in social class A and in social class D, 'qualifying' and 'not qualifying' appears to make no significant difference. The differences in the proportions reaching an occupational level 'a' or 'b' is significant for social class B and probably significant for social class C.

TABLE 4

THE PROPORTIONS OF 'QUALIFIED' AND 'UNQUALIFIED' PUPILS IN EACH SOCIAL CLASS REACHING OCCUPATIONAL LEVEL 'a' OR 'b' COMPAIRED WITH THE STANDARD ERROR OF DIFFERENCE IN THE PROPORTIONS.

Social Class.	Observed difference in proportions reaching 'a'/'b'	Standard Error of Difference
A	·022	·082
B	·454	·094
C	·234	·089
D	·080	·140

(ii) *Length of attendance at secondary school.* The results displayed in Table 5 indicate that the higher the social class the longer will the pupils remain at school.

TABLE 5

Social Classes	A	B	C	D
Mean number of years at secondary school	6·47	5·85	4·84	4·18
Difference in mean between each group and preceding one	—	·62	1·01	·66
t..............................	—	2·49	6·55	2·55

Table 6 shows that when the factor of social class is controlled the difference in the mean for 'qualified' and 'unqualified' groups are not significant in social classes A and D, though they are in classes B and C. This gives further evidence that 'qualifying' or 'not qualifying' may be of importance for the middle social classes, but of less importance in the highest and lowest social classes.

TABLE 6

THE MEAN NUMBER OF YEARS' ATTENDANCE AT SECONDARY SCHOOL OF THE 'QUALIFIED' GROUP AND THE 'UNQUALIFIED' GROUP.

Social Class	(i) Qualified	(ii) Unqualified	Difference between (i) and (ii)	*t*
A	6·71	6·30	·41	1·46
B	6·17	5·26	·91	4·9
C	5·08	4·55	·53	1·99
D	4·12	4·24	·12	·51

(iii) *Scholastic Attainments.* From Table 7 it can be seen that in both the 'qualified' group and the 'unqualified' group there is a significant relationship between social class and Senior Certificate 'O' level ($p < \cdot 01$ in each case). (The Northern Ireland Senior Certificate 'O' level, taken at the age of about 16, is the equivalent of English G.C.E. 'O' level).

TABLE 7

SCHOLASTIC ATTAINMENT SHOWN BY THE NUMBERS ATTEMPTING SENIOR CERTIFICATE 'O' LEVEL.

'Qualified Group.		
Social Classes	Not attempting Senior Certificate 'O' Level	Attempting Senior Certificate 'O' Level
A/B	14	78
C/D	54	29
'Unqualified Group.		
Social Classes	Not attempting Senior Certificate 'O' Level	Attempting Senior Certificate 'O' Level
A/B	25	35
C/D	59	17

The results shown in Table 8 show that in respect of the proportions attempting the Senior Certificate 'O' level examination, the difference between 'qualified' groups and 'unqualified' groups is not significant for social classes A and D, but is clearly significant for class B. For class C, the evidence is inconclusive.

TABLE 8

DIFFERENCES IN PROPORTIONS ATTEMPTING SENIOR CERTIFICATES 'O' LEVEL.

Class	Observed difference in proportion of 'Qualified' and 'Unqualified' Pupils attempting Senior 'O' Level	Standard Error of Difference
A	·022	·082
B	·414	·093
C	·160	·091
D	·040	·102

(iv) *Levels of Satisfaction in Present occupation.* The system of allocation to secondary schools, as it operated in 1955, does not appear to have led to any general or large scale dissatisfaction amongst the subjects over occupational status eight years later. Only 6·7 per cent. of all subjects stated that they are 'definitely dissatisfied' in their present occupations. For the 'qualified' group a significant relationship was found between social class, and the level of satisfaction in present occupation, and for the 'unqualified' group the relationship is not significant. An interesting finding was that the difference in the proportions expressing some dissatisfaction in the ' qualified ' group and in the ' unqualified ' group is significant only for social class C, the ' qualified ' being more dissatisfied than the ' unqualified.'

(v) *Participation in Further Education.* It was found that a higher proportion of ' unqualified ' subjects than of ' qualified ' subjects did not take part in further education of any kind, but the difference in the proportions is not significant. A smaller proportion of social classes C and D taken together than of A and B taken together took part in further education and the difference is significant. For those who took part in ' vocational ' further education, a distinction was made between those who followed academic courses and those who followed technical courses. There was a significant tendency for ' unqualified ' pupils and for those of lower social class to follow technical courses.

III.—Conclusions and Comment.

It is clear from the investigation that success or failure in the ' qualifying examination ' in Northern Ireland has not the decisive influence for the future that it is commonly supposed to have. On many of the criteria which make up the composite assessments of the circumstances and prospects of the subjects the influence of social class was such that ' qualifying ' or ' not qualifying ' had no significant effect. This is especially true in the extremes of the social range. Table 9 shows that fee-paying is an important factor, especially in social class A.

TABLE 9

The Distribution with Respect to Present Occupational Level, Showing Grammar School Fee-Payers and Non-Selective Secondary School Pupils Separately.

Classes	Present Occupational Level			
	a	b	c	d
A	5 : 0	14 : 0	1 : 0	0 · 0
B	3 : 0	12 : 0	10 : 15	0 : 0
C	2 : 0	8 : 0	6 : 30	1 : 4
D	0 : 0	2 : 0	1 : 15	1 : 6

Fee-payers are shown by first figure in each cell.

The Table suggests that even within the fee-paying group, social class exerts an influence. In social class A, 95 per cent. of fee-payers reach occupational, level ' a ' or ' b ' ; in class B, 60 per cent. ; in class C, 58·8 per cent. ; in class D 50 per cent.

(A larger scale investigation would be required before any statement could be made about the relative importance of the fee-paying factor and other possible factors).

In social classes B and C, the ' qualifying result ' is more important. In these classes ' qualified ' pupils tend to stay longer at school, to achieve higher academic distinction, and to reach a higher occupational level.

(*Manuscript received* 13*th January*, 1966)

THE INFLUENCE OF READING METHODS ON SPELLING

BY MARGARET L. PETERS
(*Cambridge Institute of Education*)

SUMMARY. Two experiments were carried out to test the hypothesis that different methods and different media in the teaching of reading give rise to different kinds of spelling error. Comparisons were made between children taught to read by look-and-say and by phonic methods, and between children taught in traditional orthography (t.o.) and the initial teaching alphabet (i.t.a.). When the methods are ranked for vulnerability to particular types of error, a trend is revealed in which i.t.a. produces the fewest omissions, insertions and perseverations, the look-and-say method produces the fewest doubling errors, consonant substitutions, faulty auditory perceptions and homophones. Between these two extremes, lies the phonic method which gives fewest transpositions, vowel substitutions and the greatest number of reasonable phonic alternatives. Differences between the look-and-say and phonic methods are fully consistent with the differential emphasis placed upon their 'perceptual' and 'rule-following' aspects. Differences between t.o. and i.t.a. are consistent with differences in economy of form. Breakdown of results by sex and I.Q. reveals significant differences in the lower I.Q. range. It is concluded that though overall spelling attainment is not affected, perceptual and rule-following skills influence spelling in a way which may have important implications for remedial teaching.

I.—OUTLINE OF PREVIOUS RESEARCH.

THE research grew out of the growing belief at the Cambridge Institute of Education's Remedial Centre in the need to supplement remedial reading by an intensive course in spelling if retarded children are to be effectively assimilated into normal school streams. Schonell (1942) holds that spelling attainment depends less on children's linguistic background than on school experiences, and among school experiences, the activity of reading might well be expected to be an important influence. In the early years of the century, the rejection of the phonic method as a way of teaching reading was supported in the interests of spelling. Gill (1912) attributed reversals to phonic teaching, showing that the sentence method produced spelling efficiency, at least equal to thatproduced by phonic teaching. Horn (1919) agreed that the phonic method seemed to present no advantage in learning to spell, and Hilderbrandt (1923) pointed out the danger of stressing phonetic elements rather than viewing the word as a group of significant parts or as a whole. Forty years on, however, phonics were back in favour (Beltramo, 1954), Horn (1960) noted the growing interest in the possibility that phonic instruction may contribute to spelling ability. In the Michigan research (Anderson. 1965) it is shown that the most common mis-spellings of the English and Scottish samples were more often phonetic in nature and more readable than those of the American samples at all ages. This, Ewing (1965) suggests might be attributed to the differences in emphasis in reading method on the two sides of the Atlantic. Such suggestions would seem to justify the examination of the hypothesis that different methods of teaching reading give rise to different kinds of spelling errors.

II.—METHOD OF INVESTIGATION.

The general method of the enquiry was to match pairs of children, who had learned to read using different methods and different media of teaching, and then to compare their ability in spelling.

Two separate experiments were set up.

Experiment I.

In the first experiment, sixty-nine 8-year-old children from a school using a rigorous phonic method were matched with children from a school using an equally rigorous look-and-say method. They were matched for sex, age, socio-economic level, and scores on N.F.E.R. Primary Verbal 1 and Raven's Progressive Matrices (1956).

The following tests were applied to the children :

1. (1) The Graded Spelling Test from the Standard Reading Tests (Daniels and Diack, 1958).
2. (2) A diagnostic dictation (Appendix).

In this dictation, 75 per cent. of the words occurred in Burrough's (1957) first 500 words in his study of the vocabulary of 5–6½-year-old chlidren. 96 per cent. words in the story prepared were to be found in Burrough's complete list. 86 per cent. were comparable in difficulty with words spelt correctly by children with a spelling age of 8+. Of the phonic combinations introduced 93 per cent. conformed to rule. This gave opportunity for children to use their knowledge of rules or spelling precedent, also to see where children offer reasonable phonic alternatives, or where they produce phonic alternatives not conforming to rule. There were 7 per cent. phonetically irregular words.

The diagnostic test scripts were analysed using conventional error categories—Omissions and Contractions, Insertions, Transpositions, Doubling, Substitutions, Homophones, Perseverations and Unclassified (Spache 1940, Livingston 1961), except in the case of 'substitutions,' which are broken down into consonant and vowel substitution of letters, substitution with reasonable phonic alternatives and substitution of phonic alternatives not conforming to rule. These two latter categories were included as it seemed possible that here might appear the effect of training in reading. The criterion for inclusion in the column 'Reasonable phonic alternative' was conformity with precedent. The category "errors suggesting faulty auditory perception" was added to include words incorrectly perceived or encoded.

Results of Experiment 1. Numbers of errors in each group were calculated and found to be very close in overall total, namely 1595 for look-and-say and 1550 for the phonic group. The greatest difference was between the reasonable phonic alternatives, of which the sixty-nine children in the phonic group made 149, the sixty-nine children in the look and say group 89 ($p = < \cdot 01$) (the phonic alternatives not conforming to rule, being the same for both groups) ; the phonic group made only thirty-one unclassifiable errors, as against 103 in the look-and-say group ($p = < \cdot 01$).

The results, cross-classified with respect to sex and I.Q., are shown in Table 1. Each child was classified according to the number of errors that he made in the two error categories that showed the greatest difference, viz., reasonable phonic alternatives and unclassified errors, and the resulting distributions were compared using the Kolmogorov-Smirnov two-sample test (Siegel 1956). Due to the smallness of the matched sub-samples, only one comparison exceeds the 5 per cent. level of significance. But, ***within*** the 85-115

TABLE 1

COMPARISON OF MEAN ERROR SCORES FOR TWO MATCHED SAMPLES TAUGHT BY PHONIC AND LOOK-AND-SAY METHODS.

Error Category		I.Q. (85–115) Mean No. Errors per Child			Signif. level of	I.Q. (116+) Mean No. Errors per Child			Signif. level of
		n	L. & S.	Phonic	Diff.	n	L. & S.	Phonic	Diff.
Reasonable Phonic Alternative :....	Boys	26	1·31	2·85	p< ·025	9	1·44	1·11	NS..
	Girls	25	1·32	2·40	N.S.	9	1·00	·56	N.S.
	x̄	1·31	2·63		p< ·025		1·22	·83	N.S.
Unclassified Errors	Boys	26	2·19	·35	N.S.	9	Nil	Nil	N.S.
	Girls	25	1·84	·88	N.S.	9	Nil	Nil	N.S.
	x̄		2·02	'61	N.S.		Nil	Nil	N.S.

I.Q. range, the other comparisons show consistent trends. If a gross comparison is made between all sixty-nine children in the look and say group and all sixty-nine children in the phonic group, it turns out that their results differ beyond the 1 per cent. level of significance in each error category.

For the group with the higher (116+) I.Q., it is instructive to note that differences due to different teaching methods vanish almost completely. Over both I.Q. ranges, the girls show a slight overall superiority over the boys. No data are available for children in the ' below 85 ' I.Q. range.

In addition to comparing the frequencies with which particular kinds of error occur, we can also check on the different ways in which the errors are inter-related. We can, for example, enquire whether the tendency to select reasonable phonic alternatives is associated with the tendency to select homophones. If we find that two kinds of error tend to go together, we can then ask whether the inter-correlation is higher among one group than the other. Ideally, we require some kind of factor analysis to determine the dimensionality of the underlying structure of errors. This computation has not yet been done, but some gross measures of inter-correlation have been calculated, and the following pattern emerges.

Among the look-and-say group, the highest comparative correlation is between transposition of letters and vowel substitutions. Among the phonic group, the highest correlation is between reasonable phonic alternatives and homophones. Both these correlations exceed the 2 per cent. level of significance.

Experiment II.

Having compared two methods of teaching reading and their influence on spelling ability, approaches to reading via two different alphabets were studied. 115 children learning to read in traditional orthography (t.o.) by rigorous look-and-say and phonic schemes, were compared with 115 children taught to read using the initial teaching alphabet (i.t.a.).

The look-and-say and phonic samples were compared with the i.t.a. samples with confidence since all three samples had been taught rigorously by teachers (and not only teachers but also head teachers) who were totally committed to the method used, heavily prejudiced against other methods, and whose schools were much visited as exponents of these methods. All samples were from schools in working class areas, largely council house property.

One-hundred-and-fifteen pairs were studied, matched for sex, age (within one month) socio-economic status and scores on Moray House 1A and Raven's Progressive Matrices (1956).

Results of Experiment II. Numbers of errors in each group were found to be very close in overall total, namely 3,163 for i.t.a., and 3,059 for t.o. There is a difference at $p < \cdot 001$ in the single for double letter error, and in the substitution of reasonable phonic alternatives and in homophones, indicating significantly more errors in i.t.a. than in t.o., in all three error categories the mean for i.t.a. being greater but the variance less than in t.o. The difference between t.o. and i.t.a. in faulty auditory perception, perseveration, and unclassified errors, shows significantly more errors in t.o. than in i.t.a. again $p < \cdot 001$.

As before, rough measures of association were obtained to gauge the extent to which different kinds of error clustered together. It was found that the chief cluster with respect to i.t.a., was concerned with :

(1) Substitution of reasonable phonic alternatives.
(2) Homophones.
(3) Single for double letter substitutions.

In the cases t.o., the main cluster was with respect to :

(1) Transpoition of syllables.
(2) Perseveration, and
(3) Faulty auditory perception.

A systematic breakdown of these results, by sex and I.Q., is shown in Table 2. The distributions of appropriate sub-samples were compared by the Kolmogorov-Smirov two-sample test. Once again, the girls tend to be slightly superior to the boys, and differences due to the different teaching media (i.t.a. and t.o.), are mostly reduced to non-significance in the higher I.Q. range.

The errors associated in t.o. are errors of inexactitude, viz., transposed syllables, syllables running on, syllables incorrectly perceived or incorrectly encoded, and words that are unclassifiable. i.t.a. children thus attack spelling more systematically, evidence being :

(1) Fewer unclassified errors.
(2) Fewer perseverations.
(3) Fewer transpositions of syllables.

(4) Fewer errors suggesting faulty auditory perception or encoding.

On the other hand, there are specific spelling problems, e.g., more single letters where these should be double, a spelling convention that is rational though unexpected, and this in spite of the fact that i.t.a. children have been deliberately and of necessity taught spelling rules, e.g., ' no ' instead of ' know.' Indeed, frequently the more sophisticated i.t.a. taught children fall over backwards and present ' tide ' in preference to ' tied ' which is extremely close to the i.t.a. ' tied ' !

TABLE 2

COMPARISON OF MEAN ERROR SCORES FOR TWO MATCHED SAMPLES TAUGHT BY DIFFERENT MEDIA (t.o. and i.t.a.).

Error Category		(85–115) Mean No. Errors per Child			Signif. level of Diff.	(116+) Mean No. Errors per Child			Signif. level of Diff.
		n	i.t.a.	t.o.		n	i.t.a.	t.o.	
Single for Double	Boys	33	3·24	2·03	p<·005	21	1·76	1·33	N.S.
	Girls	43	2·65	1·28	p<·005	18	·83	·56	N.S.
	$\bar{x}$		2·91	1·61	p<·001		1·33	·97	N.S.
Reasonable phonic Alternative	Boys	33	2·94	1·70	p<·005	21	1·76	1·24	N.S.
	Girls	43	2·37	1·44	p<·025	18	1·22	1·11	N.S.
	$\bar{x}$		2·62	1·55	p<·001		1·51	1·18	N.S.
Faulty Auditory perception	Boys	33	1·91	4·55	p<·05	21	·67	2·24	N.S.
	Girls	43	2·02	3·12	N.S.	18	·56	1·89	N.S.
	$\bar{x}$		1·97	3·74	p<·05		·62	2·08	p<·025
Homophones:	Boys	33	1·15	·30	p·<005	21	·19	·29	N.S.
	Girls	43	·86	·47	N.S.	18	·28	·17	N.S.
	$\bar{x}$		·99	·39	p<·01		·23	·23	N.S.
Perseveration	Boys	33	·06	·18	N.S.	21	Nil	·05	N.S.
	Girls	43	Nil	·54	N.S.	18	Nil	·06	N.S.
	$\bar{x}$		·03	·38	N.S.		Nil	·05	N.S.
Unclassified errors..........	Boys	33	·58	3·30	p<·05	21	Nil	·86	N.S.
	Girls	43	1·00	2·86	p<·025	18	Nil	Nil	N.S.
	$\bar{x}$		·82	3·05	p<·001		Nil	·46	N.S.

It would seem that i.t.a. taught children, with their more systematic and economical attack, present a more receptive base for the teaching of spelling conventions.

III.—TRENDS SUGGESTED BY EXPERIMENTS I AND II.

Livingston (1961) demonstrated similarity in pattern of errors with Masters' (1937) American investigation. Each group in the present study follows a similar pattern though the percentage of insertions is lower in every group while the percentages of omissions and unclassified spellings are higher. With these exceptions (probably due to differences in classification), the group which compares most nearly with Livingston's is the i.t.a. group, though this group shows a higher percentage of homophones than Livingston's since

Scottish children differentiate homphones orally, e.g., 'tied' and 'tide.' Indeed, it is in this particular error category that the trend from few errors in look-and-say, through phonic, to most in i.t.a. is most evident.

TABLE 3

COMPARATIVE FREQUENCY OF ERRORS FOR EACH OF THREE TEACHING SYSTEMS.

	Teaching System		
Frequency	Look-and-Say	Phonic	i.t.a.
High	6, 7, 8	11	1, 2, 3, 4, 5
Medium	9, 10, 11	1, 2, 3, 4, 5, 6, 7, 8	9, 10, 11
Low	1, 2, 3, 4, 5	9, 10	6, 7, 8

1—Doubling (Single for Double, Double for Single).
2—Substitution of consonants.
3—Non-conforming phonic alternatives.
4—Faulty auditory perceptions.
5—Homophones.
6—Omissions.
7—Insertions.
8—Perseveration.
9—Transpositions.
10—Substitution of vowels.
11—Reasonable phonic alternatives.

This trend can be seen from a study of the ranking of methods according to the percentage of errors in each category (Table 3). It will be seen that i.t.a. comes off best only in categories of error arising from lack of economy and control of output, viz., omissions, insertions and perseverations.

It will be seen that where the phonic method is superior is in producing fewer transpositions and fewer substitutions of vowels. Also when a major substitution is made by phonic children, it is more often a reasonable phonic alternative. The look and say method produces, as one would expect, much better visual attack. There are fewest doubling errors (both single for double and double for single), fewest consonant substitutions, fewest non-conforming phonic alternatives, fewest errors of faulty auditory perception and fewest homophones. It can also be seen that in all but three error categories the two methods and the initial teaching alphabet lie on a continuum from look-and-say through phonic to i.t.a. The strength of look-and-say lies in consonantal exactness, superior attempts at overall structuring, and better semantic association (e.g., in homophones where the visual aspect of the word is conjured up directly by the meaning and associations of the word). Between this and the economy and discipline of i.t.a., at the other end of the continuum, lies the phonic method, with neither fewest nor most errors in any of the above categories, yet superior in the remaining ones of fewest transpositions, fewest substitutions of vowels and the greatest number of reasonable phonic alternatives.

There appear to be important implications here for remedial teaching. For children who are finding spelling difficult, those taught originally by a look-and-say method are restricted with visual but no rational reference, those taught by phonic methods are able to make a reasonable attempt which may well be a homophone, while i.t.a. children have the sort of non-redundant 'skeletal' structure from which conventional English spellings can be readily developed.

IV.—CONCLUSION.

The use of one of two reading methods (look-and-say or phonic) or a new medium (i.t.a.), when rigorously operated, does not seem to affect the level of spelling attainment.

Differences in method or medium seem, however, to lead to differences in perception which show themselves in the type of spelling error made.

ACKNOWLEDGEMENTS.—The author acknowledges with gratitude the generous co-operation of the local authorities, heads and teachers of schools involved, in this research study, and to many students and colleagues, particularly Mr. Brian Lewis, and Mr. William Latham.

V.—REFERENCES.

ANDERSON, I. H. (1965). *Comparisons of the Reading and Spelling Achievements and Quality of Handwriting of Groups of English, Scottish and American Children.* University of Michigan.

BELTRAMO, L. (1954). *An Alphabetical Approach to the Teaching of Reading in Grade One* State Univ. of Iowa. *Dissertation Abstracts*, 14, 2290.

BURROUGHS, G. E. R. (1957). *A Study of the Vocabulary of Young Children.* Oliver and Boyd.

DANIELS, J. G., and DIACK, H (1958). *The Standard Reading Tests.* Chatto and Windus.

EWING, A. (1965). *A Scottish Viewpoint.* Appendix to Anderson, I. H. (1965) (see above).

GILL, E. J. (1912). The teaching of spelling. *J. Exp. Pedagogy*, I, 310-319.

HILDERBRANDT, E. L. (1923). The psychological anaylysis of spelling. *Pedagogical Seminary* 30.

HORN, E. (1919). *Principles of Teaching Spelling as derived from Scientific Investigation.*

HORN, E. (1906. **Spelling** *in Encl. of Educational Research.* Macmillan.

LIVINGSTON, A. A. (1961). A study of spelling errors in *Studies of Spelling.* University of London Press.

MASTERS, H. V. (1927). A study of spelling errors, Univ. of Iowa *Studies in Education*, Vo. IV, No. 4.

RAVEN, J. C. (1956). *Coloured Progressive Matrices.* H. K. Lewis.

SCHONELL, F. J. (1942). *Backwardness in the Basic Subjects.* Oliver and Boyd.

SIEGEL, S. (1956). *Non-Parametric Statistics for the Behavioural Sciences.* New York: McGraw-Hill.

SPACHE, G. (1940). Critical analysis of various methods of classifying spelling errors. *J. Educ. Psychol.*, XXXI, 2, 111-134.

(*Manuscript received*, 13*th January*, 1966)

VI.—APPENDIX.

DIAGNOSTIC DICTATION.

One day, as I was walking down Bridge Street, I heard the sound of trotting. I turned and saw behind me the shaggy dark hair of a frightened little horse. I searched in my pockets for an apple from my dinner to give him. " I know where you should be," I said. So I removed the belt of my raincoat and tied it around his neck and led him back. I opened the gate, and with satisfaction, he galloped into his own field. I was certainly very happy that now he was safe, away from the noisy and dangerous traffic.

ENVIRONMENTAL INFLUENCES ON RESPONSE TO QUESTIONNAIRES

By M. D. SHIPMAN

(*Worcester College of Education*)

SUMMARY. Investigations employing interviews or questionnaires in organisations often ignore environmental influences on individual responses. Reasons behind answers to questionnaires were investigated in interviews and group discussions. Students concealed real attitudes because of the values paramount in the college being studied.

I.—INTRODUCTION.

COLLEGES and schools are organised to produce appropriate attitudes. Although Jacob (1957), in a review of research concluded that colleges have little impact on these, he did acknowledge the influence of colleges with strong value systems. Vreeland and Bidwell (1965) have shown how the House system at Harvard enabled the Master of each House of some 400 students to exert influence on their attitudes through the affective climate. Where strong cultural forces have evolved, participants are forced to respond positively or negatively and do not remain uninvolved.

This paper examines the organisational effect on responses to questionnaires designed to measure attitudes, in a college which increased from 350 in 1961 to 700 in 1965. The hypothesis underlying it is that organisations, through their values and norms, provide an important variable in the determination of attitudes and that analysis of, and generalisations from measurements of these should only be made after accounting for this.

In 1961, an investigation was started at the Worcester College of Education, into the academic, professional and social aspects of the education of teachers and their interaction within a particular social system. Tests were given to the intake of students in this year as the start of a follow-up study. These were repeated in 1964 at the end of their course. Analysis of the first tests, however, showed that there were discrepancies between answers to questionnaires, and in interviews and group discussions. This made the reliability of the follow-up doubtful.

The reasons for the discrepancies emerged once the sociological analysis had been completed and are the subject of this paper. Many of the influences were due to the structure of the college, which was small, with a high proportion in residence. A high level of social activity involving staff as well as students had resulted in excellent personal relations between them and a tightly integrated community giving priority to maintaining consensus. This was reinforced by the universal concern with teaching as a profession.

II.—PLAN OF INVESTIGATION.

The focus on the college as an organisation undergoing rapid structural change and the position of the researcher as lecturer, determined the methods used. In 1961, a number of indices of social activity were selected and means for their continuous collection set up. Simultaneously, a programme of interviewing, questionnaires, keeping of diaries and small group discussions was designed to show how students and staff responded to and performed in this environment.

Five groups of subjects were used, students coming to the college to apply for a place, students in college, ex-students six months after leaving, college staff and staff of schools used for teaching practice. Full details are given of methods and results in Shipman (1965).

Fifteen questionnaires were used, twelve of which were designed for the investigation. Intelligence (N.I.I.P. Group Test 33) and social class were controlled for the intake in the follow-up. The Maudsley Personality Inventory (M.P.I.) and the Melvin Public Opinion Inventory (M.P.O.I.) were filled in by 93 per cent. and 92 per cent. of these 137 students at the start and finish of their course, three years later. The form of the M.P.I. used was that reproduced by Jensen (1958), containing twenty lie questions. Students were given a code number at the start of their course to use on questionnaires, enabling the answers of individuals to be correlated. The same students were frequently involved in the group discussions, interviews and questionnaires, facilitating the cross checking of results and the detection of unintended or false responses.

III.—RESULTS.

Institutional pressures influenced the frame of reference within which questionnaires were answered. The paramount concern of staff and students was with teaching and both were influenced by this in their answers. But official and informal conceptions of the teaching role did not coincide. Students maintained two levels of professional attitude, one for official use onstage, one for use backstage, out of official hearing or later on in the classroom. Inevitably questions coming from a member of staff were perceived and answered as official business.

The most marked example of this came when students were questioned about attitudes towards teaching. The table shows the results on one scale of a Lickert type schedule testing attitudes to aspects of teaching. Attitudes to restrictive discipline (low score) or free activity (high score) in the classroom.

Sample	A.M.	S.D.	Number
Students	26·3	4·0	403
Ex-Students	23·8	4·7	125
School Staff	23 4	4·5	159
College Staff	25·5	4·4	44
Total	—	—	731

These and other results could not be taken at face value. In interviews and particularly in the group discussions with ten students, tougher views were expressed. Students admitted concealing these to produce answers in line with the professional views advanced on the course. This was done because the questions were seen as official. But it was also to preserve good relations with staff.

This was also illustrated in the Lie scale incorporated into the version of the M.P.I. In 1961, forty-three out of 121 students scored over 10 on this scale and according to H. J. and S. B. G. Eysenck (1964), "... the E and particularly the N scale should be regarded with considerable scepticism." In 1964 there were twenty-three liars, seventeen of whom were also in the first group as liars. Dicken (1959) gave three attitudes contributing to lying: deliberate faking, response using an honest, but inaccurate self-assessment, and response using an idealised self assessment. Here students tended to use the last of these. Discussions revealed that questionnaires were treated as examinations in education theory. There was no difficulty in holding simultaneously right beliefs and applicable ones, official and unofficial, formal and informal, backstage and onstage. This even persisted when students had completed their work in college and were about to leave. Once they had left and become teachers the cultural influence was absent and the tougher attitudes emerged.

This concept of impression management, the use of a veneer, developed by Goffman (1958) applies whenever persons of different status interact. It is particularly relevant when one party is subject to assessment by the other. If this assessment is formalised in examination or questionnaire, these situations will produce distorted answers.

One result of the consensus of the college was to minimise non-response. The lowest response was 85 per cent., even the postal questionnaires from ex-students achieving 91 per cent. However, throughout the study a minority of 10 to 15 per cent. showed persistent resentment of the inclusive, constraining nature of the college and this was reflected in their response to this research. Included in this group were 6 per cent. who were non-responders through design rather than accident. They were all male in a college with a majority of women. While objecting to constraints they were still affected by them, reacting by activity rather than withdrawal. They were socially active and bent on reforming the system, not forming a sub or contra-culture. This concern with the college led to their filling up any schedules relevant to the working of the college as an organisation, but they rejected the material on personality, attitudes and intelligence. Consensus meant that the majority responded well to requests for personal information, even though code numbers did not guarantee anonymity. It was the most interesting, non-conforming group who were not represented.

IV.—DISCUSSION.

These results show that where the culture and social structure of an organisation is ignored, while participants in it are being questioned, there is a danger that extraneous variables will intervene. Further, these may be operative at many levels. In some research, the effect of cultural pressures may be under review, but frequently results are expressed in terms of students or pupils or workers without this reference. Its importance increases as the organisation studied approaches a 'total' institution in the sense used by Goffman (1961). This college possessed a pervasive culture and satisfield most individual needs on the premises. Yet it was one of the largest in the country and smaller colleges and schools would be likely to exert even greater influence.

The danger of ignoring these environmental influences is not only that 'lying' can not be measured, but that it may not even be detected. In the example here a false, idealised professional image was used. Interviews with 125 applicants for places showed that this was part of the process of identifying with teaching as a career. It was reinforced after arrival at college by the commitment to teaching and the pressures to preserve consensus. But equally common would be lying to fit other images favoured by the informal, peer culture. Such a danger would seem to be present among younger children and adolescents where sub and contra-cultures are liable to be as important a reference as the overall culture of the school. Where comparisons across time are made the dangers increase, for not only will the institution have changed, but also the informal pressures.

Even at the technical level of research an appreciation of the organisational effect seems essential. Non-response is often difficult to calculate quantitatively when an outsider visits the institution. Estimations of the characteristics of non-responders are rare. Postal questionnaires, where a minority only often respond are probably sampling not only the ultra responsive but the ultra conformist. To erect any statistical calculations on such a basis is hazardous.

The need for such caution was apparent from the number of researches which ignore these factors. The intake studied here for the follow-up filled in four questionnaires during their three years in college, apart from taking the brunt of this study. Each was long and asked for much data readily available in college files. None of them was accompanied by any attempt to study the social structure of the college. In three of these cases the researcher never came personally, the questionnaire being delivered by post. However carefully this data was processed, its reliability could never be estimated and each time the students filled one in, the chances of the next being treated seriously were reduced.

Finally, examination of the cultural context in which data was obtained would act as a check on easy generalisation from research in a particular institution. If results reflect the social system, then generalisation may be impossible. The teacher who rejects research results because they would not apply in his school is expressing this view that different social systems induce different responses to similar stimuli.

In conclusion, research strategy should include :

(1) Investigation of institutional pressures to help in the analysis of results.

(2) Lie detectors or a variety of techniques to facilitate cross checking of results (Gibson 1962 and 1964).

(3) A multi-disciplinary theoretical framework, whereby the sociological influences on psychological data would not go undetected.

V.—REFERENCES.

DICKEN, C. F. (1959). Simulated patterns on the Edwards' personal preference schedule. *Applied Psychol.*, 43, 372-377.

EYSENCK, H. J. and S. B. G. (1964). *The Manual of the Eysenck Personality Inventory.* London : Univ. of London Press.

GIBSON, H. B. (1962). The lie scale of the Maudsley personality inventory. *Acta Psychol.*, 20, No. 1, 18-23.

GIBSON, H. B. (1964). A lie scale for the junior Maudsley personality inventory. *Brit. J. Educ. Psychol.*, 34, 120-124.

GOFFMAN, E. (1958). *The Presentation of Self in Everyday Life.* Edinburgh : Univ. of Edinburgh Soc. Sci. Res. Centre.

GOFFMAN, E. (1961). *Asylums.* New York : Anchor.

JACOB, P. E. (1957). *Changing Values in College.* New York : Harper Bros.

JENSON, A. R. (1953). The Maudsley personality inventory. *Acta. Psychol.*, 14, 314-325.

SHIPMAN, M. D. (1965). *Personal and Social Influences on the Work of a Teacher Training College.* Ph.D. thesis, Univ. of London.

VREELAND, R., and BIDWELL, C. (1900). Organisational effects on student attitudes : A study of the Harvard Houses. *Soc. of Educ.*, 38, No. 3, 233-250.

(*Manuscript received* 8*th March*, 1966.)

VARIABLES AFFECTING AESTHETIC APPRECIATION, IN RELATION TO AGE

By E. E. RUMP and VERA SOUTHGATE

(*Department of Psychology and School of Education, respectively, University of Manchester*)

Summary. The pictorial interests and preferences of 7, 11 and 15-year-old children and of adults were investigated. Stimulus material consisted of seventy-six varied items displayed in an art gallery. Some concordance was found, using three measures of appreciation. The 7 and 11-year-olds strongly preferred pictures realistically depicting familiar objects. Brightly coloured items were preferred by all age-groups. Abstract paintings preferred by some older subjects were liked for their colouring. Comments mainly referred to objects represented, or were expressions of approval; 7-year-olds made some egocentric remarks. Sex differences, especially marked for younger children, were found for some pictures. The stated preferences of an interviewer had a positive suggestibility effect, whereas the purported preference of a teacher had a negative effect on children's preferences.

I.—Introduction.

Relatively little has been written about the role of art *appreciation* in fostering perceptual awareness and aesthetic enjoyment in children. However, the ' pictures in schools ' movement has been growing, and in 1963, an exhibition was organised by Manchester University School of Education of pictures selected, as suitable for hanging in schools, from the Manchester Education Committee's loan collection and the Whitworth Gallery collection.

The exhibition included thirteen original oil paintings and six paintings in other materials by contemporary artists of repute, twenty-one original prints, high quality reproductions of eight 15th to 18th century works and of ten 19th and 20th century works (from Giovanni Bellini to Picasso), eleven paintings by local pupils, and seven textiles by contemporary designers. These seventy-six items were framed, and hung in a well illuminated gallery.

The interests and preferences of pupils visiting the exhibition were investigated. There have been several relevant studies of preferences, but many have used few items (e.g., Calkins, 1900), or only small reproductions or designs (e.g., French, 1952), or merely diagrammatic material (e.g., Berlyne, 1958). The main advantage of the present study was, then, that a wide variety of suitable items of aesthetic value was available.

No attempt was made to measure the development of good taste (e.g., Littlehohns and Nedham, 1933); it is difficult to equate this with ' concensus of opinion ' especially if derived from laymen as, for example, in Eysenck's studies (1940, 1941b). Rather, the aim was to investigate particular features of developmental change. The general hypothesis was that preferences of the younger children would be strongly influenced by simple stimulus characteristics, whereas older children would be more affected by such factors as composition and style.

II.—Method.

It was arranged that four parties of pupils visiting the exhibition should take part in the investigation. There were twenty boys and twenty-one girls aged 7, fourteen boys and twenty girls aged 11, and twenty-eight boys and twenty-one girls from two technical high schools, aged 15. In addition, seven male and eleven female teachers attending a conference associated with the exhibition also participated. The older subjects probably have above average artistic interest, and the effects of age and of developing interest are, therefore, not absolutely separable in this study. It is, nevertheless, felt that the results obtained do add to the understanding of factors affecting aesthetic preferences in relation to age.

(a) *Treatment of subjects*

Interviewers who conducted the children round the exhibition were five men and five women, graduates in Psychology or with qualifications in Education. They were given general instructions, received notes on methods to be used, and a final briefing before each session. The importance of being friendly, but entirely neutral with respect to the pictures, was emphasised.

The children's appreciation of the pictures was investigated by three different methods progressing in specificity. In stage 1, the children were escorted round the exhibition in groups of two to five (usually all boys or all girls); there were eight such groups at each age level. Half the groups toured one way and half the other way round the room. During this period the interviewer's task was to ensure that all pictures were seen by each child in his group, and to record as much as possible of the 'free comments ' made about the pictures. The interviewer said as little as possible, though answering such queries as titles and painters' names. After all the pictures had been seen, the interviewer recorded which paintings seemed to interest the group most. The children were then allowed to ' explore ' other sections of the Gallery, or in the case of the 7-year-olds were given drawing materials.

In stage 2, the 7 and 11-year-olds were escorted individually back into the exhibition room. They were asked " Which picture do you like best ? " and then " What is it about the picture that you like ? " They were asked the same questions concerning their second choice. During this period, the children were allowed as much time as they wished to look round the exhibition again. The 15-year-old pupils and the teachers were each given a corresponding questionnaire to answer after seeing the exhibition.

In stage 3, about a quarter of the children (randomly selected), and the teachers, were studied further by use of the paired comparison method. The children were interviewed individually, and the teachers continued the questionnaire. The first specific hypothesis tested in this way was that children, and particularly younger ones, would prefer representational paintings to more abstract paintings. Four pairs of paintings were used, one of each pair being representational and the other fairly abstract in style but similar in size and colouring. The subject was asked " Will you tell me all you can about this painting ? " for both, then " Which of the two pictures do you like better ? " and " What makes you like this one better ? " Order of presentation was systematically vaired in this and the following tests.

The second hypothesis was that children, and particularly younger ones, would prefer pictures with a familiar content. Pairs of pictures were selected, so that subjects might choose between one with a familiar content and one with an unfamiliar content, but similar in size and colouring. Unfortunately, an independent judge subsqeuently agreed with this selection for only one such pair, and results are reported only for this pair.

Two other specific hypotheses were tested with the children only. It was considered that correct interpretation of perspective would increase with age.

Two close but not overlapping objects in a picture were indicated, and the subject asked " Which looks nearer, . . . or . . . ? " This was repeated with another picture, with ' further away ' replacing ' nearer.' Four different pictures were used. The final specific hypothesis was that prestige suggestibility would decrease with age. This was tested by indicating two fairly similar pictures and saying : " Your teacher (*or* ' *I* ') quite like/s this one. Which do you like better ? " Four pairs were used, and the item suggested was varied systematically.

(*b*) *Classification of pictures.*

Because of the large number of varied items in the exhibition, it was possible to investigate the influence of general characteristics on appreciation by classifying items in different ways. The first was in terms of the sections in which the exhibition was arranged, as listed above. Other characteristics were determined by the ratings of two judges. These were the first author and a colleague, working independently. Each picture was rated as representational or abstract, of familiar or unfamiliar subject matter for children, simple or complex in design, and brightly coloured (with primary colours of high saturation) or dull (with subdued, uniform, or monochromatic colouring). In each case a third ' uncertain ' category was allowed. The proportion of pictures for which the judges were in agreement ranged from 55 per cent. to 75 per cent. for the four categories. In about 25 per cent. of the cases, one of the judges used the ' uncertain ' category, so that the proportion of complete disagreements was quite small, ranging from 0 per cent. for the representational-abstract classification to 13 per cent. for the simple-complex classification. This degree of reliability is sufficient for the classifications to be used in the analysis of preferences. Pictures classified by both judges as ' uncertain ' or in opposite categories were excluded ; however, for the representational/abstract variable, such ambiguous items were counted as ' abstract ' (since the number of pure abstracts was very small).

These binary classifications resulted in the following proportions : 71 per cent. representational, 69 per cent. familiar in content, 62 per cent. simple in design, and 67 per cent. brightly coloured. This suggests that these characteristics were deliberately selected for the exhibition on the assumption that they would appeal to children. One of the aims of the investigation was to test the validity of this assumption.

The two judges also noted any paintings which they considered would appeal by their content especially to boys or to girls Very few pictures were selected, and a third judge was asked to pick out such paintings. 20 per cent. of the paintings were then considered to appeal to boys by two or three judges, and 12 per cent. to girls. The extent of sex differences in preferences for these pictures could thus be studied.

III.—RESULTS AND DISCUSSION.

(*a*) ' *Free Comments* ' *about the pictures.*

The comments children made during stage 1 proved difficult to categorise. The analysis was refined in several stages, but even so discrepancies in categorising by two judges amounted to 7–28 per cent. of comments for the three age-classes. Therefore, only overall mean percentages of 685 comments are given and main age-trends (Table 1), using the first author's ' arbitrated ' categorisations in cases of disagreement between judges.

Most comments are simple lists of represented objects (37 per cent.), e.g., " That's Mary and Jesus of Nazareth," by a 7-year-old (cf. Vernon, 1940) ; or evaluations (36 per cent. approving, 9 per cent. disapproving), e.g., " I'd buy that ! " by a 15-year-old. Several comments on such features as technique and composition are recorded (11 per cent.), e.g., " If it was cut in half, you wouldn't notice " by an 11-year-old ; and on degree of realism (4 per cent.), e.g., " You couldn't get his legs turning that way " by a 15-year-old. The youngest age-class made few such comments, and indeed, some of their remarks were purely personal, e.g., " I went to Blackpool for a week," said a 7-year-old on seeing a painting of *Eastbourne Beach* (cf. Dewar, 1938).

TABLE 1

" FREE COMMENTS " ABOUT THE PICTURES : DIFFERENCES BETWEEN AGE-CLASSES.

Category of Comment	Nature of Age-Class Differences	Relevant Statistic	p (1 tail) less than
Disapproval	+ ve correlation with age	Tau = + 0·41 ..	0·01
Evaluation of colour	+ ve correlation with age	Tau = + 0·29 ..	0·05
On composition, harmony, etc.	More frequent with older children than with 7-year-olds	Fisher-Yates Test	0·05
On realism or perspective	More frequent with older children than with 7-year-olds.	Fisher-Yates Test	0·05
Merely ' personal ' reference	—ve correlation with age	Tau = —0·40	0·05

NOTE .
The Fisher-Yates exact test for 2 × 2 contingency tables was applied by considering the number of groups which did, and did not, make such comments, for the 7-year-olds and for the older children. The calculation of Tau was based on the percentages for independent viewing-groups, and its level of probability allowed for ties.

(*b*) *Concordance* : *the degree of agreement within age classes.*

Three measures of appreciation are available, each of which may be related to individual pictures : ' approving comments ' during stage 1, recorded ' interests ' of the groups, and individually expressed first and second preferences.

The extent of the observed agreement in preferences amongst children of each age-class must be measured against the degree of coincidence to be expected by chance. In the case of individual preferences, since each person made one (first) choice out of seventy-six possibilities, the chance expected frequencies of pictures preferred by none, by one, by two, or by three or more subjects may be calculated on the basis of the Poisson distribution. These may then be converted to give the frequencies of *subjects* expected to choose the same picture as no other subject, as one other subject, or as two or more to her subjects (on the basis of the null hypothesis that each item had an equal probability of being chosen). In the case of ' approving comments ' and ' interests ,' the calculation of chance expected frequencies is not as simple, since the number of observations is not identifical for each group. The exact probabilities may be calculated, however, on the basis of the numbers of observations recorded for each group These too may then be converted to the expected number of *groups* approving the same item as no other group, as one other group, or as two or more other groups.

TABLE 2

Example of a Frequency Distribution, Showing the Above Chance Agreement Within the Age-Class.
(Individual Preference of 7-year-olds, N=41)

No. of Subjects preferring a picture chosen by :	Observed Frequency		Expected Frequency*
	First Choice	Second Choice	
No other subject	13	18	24
1 other subject	6	10	13
2 or more other subjects	22	13	4

* Given to nearest whole number, but calculated to 5 significant figures for χ^2 values in Table 3.

A typical example of the resulting observed and expected distributions is given in Table 2, for individual preferences of 7-year-olds. It is clear that the number of subjects choosing pictures which are also preferred by two or more other subjects, is much greater for both first and for second choices than would be expected by chance. All the obtained distributions show this type of deviation from chance. The extent of the concordance within an age-class is indicated by the deviation of observed from expected frequencies, and measured by χ^2. (Unfortunately, expected frequencies for the ' two or more subjects ' category, are rather less than 5, but no better evaluation is available, Yates' continuity correction being applied.) The results are shown in Table 3 : the χ^2 values given for the children's individual preferences are averages of the values for first and second choices.

TABLE 3

The Degree of Concordance Within Age-Classes, Indexed by Chi-Square Values.

Age-class	Individual Preferences			Approving Comments			Interests Shown		
	No. of Ss	χ^2 (2 d f.)	$P<$	No. of Groups	χ^2 (2 d f.)	$P<$	No. of Groups	χ^2 (2 d.f.)	$P<$
7-year-olds ..	41	49·3	·01	7*	17·7	·01	8	15·1	·01
11-year-olds	34	67·5	·01	8	51·0	·01	7*	8·1	·05
15-year-olds	49	64·2	·01	8	11·5	·01	8	9·2	·01
Adults	18	6·3	·05						

*The tabulated number is less than the full number in the study, because a few observations were missed for administrative reasons.

There is a significant degree of concordance in all cases. Some items tend to be appreciated, and others not, by each age-class. However, it is also true that the extent of this agreement amongst subjects is small. When faced with a wide range of pictures of good quality, the children are, to a large but not complete degree, independent in their preferences ; thus, 38 per cent. of the children gave as their first choice a picture selected by nobody else in their classes. And sixty-eight of the seventy-six items received at least one ' approving comment.' There are no consistent age trends in the extent of concordance.

“Intelligence Assessment: A Theoretical and Experimental Approach”

By

H. J. Eysenck

British Journal of Educational Psychology. Vol. 37. February, 1967.

Copies of this important article are to be printed separately and will be held for sale until June 30th, 1967.

Off Prints will be available as follows :—

Single copies (up to 10) including postage			1s. 0. each.
Batches of 10 copies	,,		7s. 0.
Batches of 100 copies	,,	£3	0s. 0.

Orders with remittance should be sent to the Editor :—

Mr. L. B. Birch,
Institute of Education,
The University,
Sheffield, 10.

It is very much regretted that we do not have the facilities to send out bills and invoices and hence cannot accept orders without the money.

However, the teachers show a very much smaller concordance than the children. No picture was most preferred by more than two teachers, but with all classes of pupils at least one picture was chosen by four or more children.

Eysenck (1940) found a general (popularity) factor, when artists and others gave preference ranks for a wide variety of materials. But the average corrected correlation with the mean preference ranking was only 0·34, and this hardly " accounts for what we call ' beauty ' in pictures " as Eysenck supposed. McGlynn and Dearnaley (1963) found that adults who had similar art training tended to agree in their assessments of paintings of an abstract artist, but untrained adults showed no common preferences. On the other hand, Gordon (1955) found greater consistency of preferences for oil paintings amongst uninterested laymen than amongst an assortment of ' art experts ' whose training varied. Concordance thus appears to be greater with a wide variation amongst stimuli, and with subjects having a similar art training.

Few studies have compared concordance at different age levels, and views vary. Berry (1942) states " pictures of animals appear to draw the greater number of favourites," which implies some concordance amongst children. However, Dashiell (1917) found no consistent preferences amongst kindergarten children for simple colour combinations ; and Winstanley (1963) found no common preferences when children choose pictures for their class-rooms. The evidence suggests, then, that if children's preferences are found for pictures which are of good quality and suitable in content, common preferences do occur, but only to a small extent at each age-level.

(c) *The degree of agreement between age-classes.*

To what extent do different age-levels concur in their preferences ? To answer this question, the method used is straightforward, though not elegant. The preferences stated by one age-class are dichotomised into those which coincide with at least one choice by another age-class, and those which do not. The ' chance ' proportion of coincidences is the number of different pictures chosen by the second age-class divided by the number of items. χ^2 values may be calculated from the differences between these observed and expected values. A corresponding value may be derived by considering the number of choices made by the second age-class which coincide with the first age-class choices (The two χ^2 values are not identical, being analogous to regression coefficients). These two values may be averaged to give a measure of the degree of agreement between the two age-classes. This analysis was completed for all age-class comparisons, for each of the three measures of appreciation (see Table 4).

The closer the age-classes compared, the greater the similarity in their preferences as indicated by the χ^2 values. Clear agreement occurs only between adjacent age-classes. The 7-year-olds show no agreement with teachers in their preferences. With regard to particular items, John Powell's *Fairground at Arundel* is very popular with both 7 and 11-year-olds, and Canaletto's *Quay at the Piazetta* (reproduction) is popular with both 11 and 15-year-olds. Reproductions of Renoir's *On the Terrace* and Degas's *Dancer in her Dressing Room* are liked by 11-year-old girls, and Maratta's *Holy Night* by 15-year-old girls. The teachers have no clear popular choice, but complex lithographs such as an abstract by John Piper are preferred by four teachers. Representative of the items not appreciated by any subject is a tapestry *Adam* designed by Keith Vaughan : a typical comment is " Too dull. I can't understand it."

Young children do not always prefer different items from adults. Indeed, Eysenck (1942) states that " laymen and even young children tend to agree on

TABLE 4

THE DEGREE OF AGREEMENT BETWEEN AGE-CLASSES, INDEXED BY CHI-SQUARE VALUES.

Age-classes compared	Age-class 'step-interval'	Individual Preferences		Approving Comments		Interests Shown	
		Mean χ^2 (1 d.f.)	P<	Mean χ^2 (1 d.f.)	P<	Mean χ^2 (1 d.f.)	P<
7 and 11-year-olds ..	1	24·1	·01	0·8	n.s.	1·3	n.s.
11 and 15-year-olds ..	1	41·2	·01	4·7	·05	13·6	·01
15-year-olds and Adults	1	7·6	·01				
7 and 15-year-olds ..	2	2·8	n.s.	1·1	n.s.	3·7	n s.
11-year-olds and Adults	2	4·5	·05				
7-year-olds and Adults	3	1·2*	n.s.				

* This deviation is in the direction of *less* agreement than expected by chance.

the average with the judgment of experts." Margaret Bulley (1934) found that children aged 7 and above from higher social class homes or at good schools scored as highly as teachers on her test of 'good taste.' And Burt (reported in Valentine, 1962), using postcard reproductions found that the preferences of 12-17-year-old grammar school children correlated 0·6—0·7 with experts', and even 6-11-year-olds gave correlations with experts of 0·4—0·5.

(*d*) *Stated Reasons for Choice.*

In analysing the reasons given for preferring a picture, two judges inspected the scripts and decided on suitable categories corresponding to some extent with those used in analysing 'free comments.' One judge categorised all reasons given. Small categories were then combined, and the analysis repeated by two judges using only the most emphasised reason for each choice so that subjects would have equal weight. Disagreement between judges occurred in only 10/192 cases.

TABLE 5

STATED REASONS FOR PREFERENCES.
(1st and 2nd choices combined)

Category	Percentage Reasons for Each Age-Class				Reliability of Age-Class Differences:	
	7-year-olds	11-year-olds	15-year-olds	Adults	χ^2 (3 d.f.)	P<
	%	%	%	%		
Realistic, particular objects, detail....	71	60	45	17	9·5	·05
Colour, colour harmony	19	17	27	44	4·9	n.s.
Composition, harmony, technique, atmosphere	4	18	24	33	12·4	·01
Other	6	5	4	6		

Percentages in each category are given in Table 5. For 7-year olds, the main reason for liking a picture concerns the objects represented. The proportion of simple 'realistic' reasons decreases significantly with increasing age. Several previous investigators have found that young children up to about 9 years, when asked reasons for preferences, simply name or describe objects depicted (Calkins, 1900; Hildreth, 1936; Valentine, 1962). With the older children, this type of reason is sometimes quite elaborate: "It is realistic I like the expression. The hands are good. I like the way the fingers are holding the pipe" (11-year-old). Teachers may be even more sophisticated in their descriptions: "It is a picture to make one think about the intelligence and beauty of animals."

Colour is the next most frequent category, and it is to be noted that it does not decrease with age. (This result is confirmed and discussed in detail in a later section.) "The colours are ready to dance" said an 11-year-old girl about *Dancer in her Dressing Room*. Valentine (1962) found that colour was ranked on average the most important factor. He also found, as did Littlejohns and Needham (1933), that 'composition, technique, etc.' are infrequently given by the younger children. However, in this study such reasons do occur with 11-year-olds: e.g., a landscape is liked because of "How it's set out. It's realistic; two people, like they've just packed up their tent to go. The forest is in the right proportion. The light and shade, and the way the bird is perched on the wires." Whorley (in Meier, 1933) found that even if young children do not talk about composition, they may show a sense of compositional balance on tests using suitable play materials.

The 'feeling' of a picture is mentioned by several older subjects, particularly 15-year-old girls in relation to reproductions of early oils: e.g., *Holy Night* is liked because "It is peaceful." A teacher likes Tintoretto's *Christ at the Sea of Galilee* because of "the qualities of restlessness, portents of doom or disaster, deep emotion, human relationships . . ." Valentine notes that such 'atmosphere' responses occur only with older subjects, and this is also reported by Dewar (1938). Such responses amount to 11 per cent. of the reasons given by 15-year-olds and teachers combined. Pickford (1948) termed emotional feeling in pictures, 'The Aesthetic Factor'; it was obtained as the main factor in his study merely because of his choice of criteria used for rating. The factor had only a small correlation with 'liking.'

A few older subjects like a picture because of its sybmolism: e.g., an 11-year-old girl likes *Holy Night* because "All the faces seem to light up from the baby. You know—they say 'The Light of the World'." A few teachers comment on originality. Younger children sometimes give purely personal reasons, like "I'm going to be a sailor, I like the suits of sailors."

(*e*) *Effects of various characteristics on appreciation of pictures.*

(i) *Type of Exhibit.* The relative appreciation of different types of picture is shown in Table 6. In this and the following sections, only first choices are stated for individual preferences, although differences and trends are indicated only if shown by second choices also. The figures quoted in Table 6 and below are given as the number of appreciations for a type of picture, given per item and per observation × 1,000 for ease of comparison.

Pictures most popular with the children are reproductions of early paintings and recent original oils. The least popular are framed textiles and original prints. These preferences may be partly determined by the chracteristics discussed below. Thus, the early reproductions are all representational, and

TABLE 6

APPRECIATION FOR PICTURES OF DIFFERENT TYPES, ALL CHILDREN COMBINED.

(No. of Preferences, etc./picture/1,000)

Measure of Appreciation	Type of Picture (in approximate order of preference)						
	C15th—C18th Reproductions	Original Oils	C19th—C20th Reproductions	Pupils' Paintings	Original Water Colours	Original Prints	Framed Textiles
Individual Preferences..	27	20	15	16	14	5	1
Approving Comments ..	20	18	14	11	13	11	5
Interests Shown	32	18	15	11	8	8	4

NOTE:

The total number of observations, and the deviation from the expected frequency distribution are for the three measures respectively: 123 ($\chi^2=48{\cdot}3$, 6 d.f., $p<{\cdot}01$); 224 (χ^2)$=22{\cdot}2$, 6 d.f., $p<{\cdot}01$); and 103 ($\chi^2=34{\cdot}2$ 6 d.f., $p<{\cdot}01$). The expected frequencies are based on the null hypothesis that the probabality of 'being appreciated' is equal for all items.

most are familiar in content (87 per cent.) and brightly coloured (also 87 per cent.); and these characteristics are found to be highly asscoiated with popularity. On the other hand, the textiles include a relatively high proportion 'dull' in colour (71 per cent.), and with unfamiliar content (29 per cent., compared with the overall 13 per cent.).

The following significant age differences are found. The χ^2 values are derived from the raw frequencies, using marginal totals as the basis for expected frequencies. The number of choices for recent oils is thirty for 7-year-olds, and only fifteen for older children ($\chi^2=4{\cdot}0$, 1 d.f., $p<{\cdot}95$). Conversely, the number for pre-19th century reproductions is forty-four for 15-year-olds, and only seventeen for younger children ($\chi^2=6{\cdot}5$, 1 d.f., $p<{\cdot}05$). Similar differences are found with regard to the interests shown. Significantly, more adults than children prefer prints ($Xf=7{\cdot}3$, 1 d.f., p—·01). These differences arise partly because of the particular subject matter of the most popular pictures, given in section (c) above.

Preferences for different tyeps of pictures have been mentioned in several previous studies comparing two or three types. Reproductions of 'Old Masters,' are frequently preferred (Valentine, 1962; Littlejohns and Needham, 1933). Differences in the present study are not as extreme as in these previous reports perhaps because all items had been carefully selected. Some Impressionist reproductions are very popular with 11-year-olds because of their bright colours and subject matter; Eysenck (1941a) found that liking Impressionist paintings is associated with youth.

(ii) *Representational versus Abstract Pictures.* The relative appreciation of representational as against abstract (or 'semi-abstract') pictures was analysed. However, since there are only two categories, the relative preferences may be conveniently given as the ratio of preferences for representational items to those for abstract items, corrected for the ratio of such items in the exhibition.

These preference ratios are given in Table 7. Where ratios are similar for different age-classes, overall values are quoted.

TABLE 7

RELATIVE APPRECIATION OF PICTURES FOR THREE CHARACTERISTICS.

Measure of Aprpeciation	Subjects	Preference Ratios (No bias gives P.R. = 1)					
		R'presentational/ Abstract	χ^2	familiar Un-familiar	χ^2	Simple/ Complex	Xf
Individual Preferences	1 and 11-yr.-olds	7·9	34·6‡	11·5	9·4‡	2·7	7·7‡
	15-year-olds			1·5	n.s.	0·8	n.s.
	Adults	0·8	n.s.	0·8	n.s.	0·3	* †
Approving Comments	7 and 11-yr.-olds	1·9	12·3‡	2·3	6·3*	1·4	n.s.
	15-year-olds. ..			1·3	n.s.		
Interests .... Shown	7 and 11-yr.-olds	2·7	11·8‡	1·7	n.s.	2·1	5·0*
	15-year-olds					0·9	n.s.

χ^2 values, with 1 d.f., are for deviations from the expected frequencies based on the proportions of items in each category.

* $p < \cdot 05$. † Vinomial test used. ‡ $P < \cdot 01$.

The children strongly prefer representational pictures. This is so for all age-classes and measures of appreciation. The bias is greatest for individual preferences whilst ' approving comments 'are somewhat more evenly spread over the items. There are no reliable age-class differences, except that the teachers do not show any bias against abstracts ; teachers and children differ significantly in this respect ($\chi^2 = 12 \cdot 9$, 1 d.f., $p < \cdot 01$). A somewhat more marked age-trend is found with the paired-comparison results. The proportions choosing the more representational painting are, in order of increasing age-level, 90 per cent., 89 per cent., 67 per cent., and 53 per cent. The difference between the two younger and two older age-classes combined is significant ($\chi^2 = 3 \cdot 85$, 1 d.f., $p < \cdot 05$). It is to be noted, however, that 7 and 11-year-olds do not differ, and this is the case for most of the following analyses.

Representational paintings are liked mainly for their ' realism ' and ' ease of understanding ' (67 per cent. of reasons, compared with 0 per cent. for abstracts) ; whilst abstracts are liked mainly for the colouring (60 per cent., compared with 8 per cent. for representational pictures). The contrast is significant using Fisher's exact test, $p < \cdot 01$. Results are similar if reasons for preferences given in stage 2 are analysed separately for abstract and representational pictures. The only 7-year-old to prefer an abstract gives an unusual reason : " It's got a long frame." Some of the comments about abstracts are caustic : " ridiculous " (15-year-old), " a big mess " (11-year-old), and " I don't like old junk " (7-year-old) ! As in Hildreth's study (1944) of children's reproductions of designs, a search for meaning is apparent : abstracts are described as " like someone's bedroom who doesn't keep it tidy " (10-year-old) and " a policeman getting ready to go out " (7-year-old). Older subjects preferring abstracts see them in a different light. Apart from colouring, the other main reasons for liking abstracts are " excitement, originality, or feeling of wonder " (20 per cent.) and ' composition ' (13 per cent.) ; percentages are only 8 and 6 per cent., respectively, for representational pictures.

Most previous research emphasises a general dislike of abstract art. Thus, Pickford (1948) found with students that 'liking' and 'representational accuracy' were strongly related in his bipolar factor. Valentine (1962) even wrote of "the lack of real beauty found in the unconventional pictures." But both Peel (1945) and Gordon (1955) found that, whilst 'laymen' usually prefer representational pictures for their realism, 'artistic' subjects tend to prefer abstracts for their colour and composition. And Tomlinson (1947) considered that "children, when given the opportunity to choose their own pictures for their classrooms, invariably choose modern examples, work of the post-impressionists in particular."

The evidence shows both extremes to be wrong. Children's drawings and preferences begin by having "a more cognitive than aesthetic meaning" (Harris, 1963). Even the very young child attempts representation in his drawings, although 'symbolic' rather than 'realistic' (Arnheim, 1956). Understanding and appreciating the complex symbolism and aesthetic qualities of abstracts appears to develop during adolescence.

(iii) *Understanding perspective.* If young children so clearly prefer representation pictures, can they correctly interpret the perspective involved? The results of the perspective test are clear: thirty-five children made two perspective judgments each, and the overall proportion correct is 89 per cent. There is no age-trend; even the majority of 7-year-olds can thus interpret pictorial cues like relative size and position.

Children at first represent space by simple ground lines in their drawing (e.g., Barnhart, 1942). But this does not imply that they cannot perceive perspective. Wohlwill (in Lipsitt and Spiker, 1963) found that although it was only by the age of 9 years that more than half the children represented perspective in drawings, even 7-year-olds were clearly affected in their perceptual judgments by a depicted perspective grid.

(iv) *Familiar content versus unfamiliar content.* In analysing this characteristic, the 'ambiguous' items were not included (see Table 7). The 7 and 11-year-olds strongly prefer pictures with familiar content, though their 'comments' and 'interests' are not so severely restricted. The 15-year-olds have significantly fewer choices for 'familiar' items than the younger children ($p < \cdot 05$, using Fisher's exact test with Tocher's modification). Teachers show no bias due to familiarity of content. Age differences are also found in the paired-comparison test: none of the eight teachers and 15-year-olds tested chose the picture of a fish in preference to one of an 'enchanted forest,' whereas three of four 7-year-olds preferred the familiar fish ($p < \cdot 05$, Fisher's test).

Similar results have been reported. For example, Littlejohns and Needham found that most children under 11 years liked a picture depicting horses, hens and a child, whereas older subjects described it as crude or gaudy. Similarly, Dewer (1938) noted that 7-year-olds prefer pictures of familiar subject matter. Of course, even young children like some novelty if it is within a familiar context, such as new toys (Cantor, in Lipsitt and Spiker, 1963). Children may thus be led from the relative sterility of comics to the creative novelty of, say, some Polish cartoon films.

(v) *Simple versus complex pictures.* 7 and 11-year-olds show a preference for simple items (i.e., with few 'elements'), whilst no such bias is shown by 15-year-olds (see Table 7). The difference is significant for individual preferences ($\chi^2 = 4 \cdot 5$ 1 d.f., $p < \cdot 05$). The teachers show a significant preference for complex items. As Harris (1963) remarks, the child "moved from simple to complex forms, (and this) extends also to his preferences in art."

Similar results have been obtained by others, mainly using diagrammatic material. French (1952) found that most first-grade children preferred simpler items, whereas teachers consistently preferred more complex ones. Spitz and Hoats (1961) found a tendency to select less complex figures amongst 8 and 16-year-olds; whilst Berlyne (1958), using similar meterial, found that adults tended to prefer the more complex figures. Eysenck (1941b) found that some adults preferred " regular polygons with few sides " whilst others preferred " not so regular, many sided polygons." Clearly Eysenck's formula " aesthetic pleasure=order=complexity " may be " too general to be useful," as he admits (1942).

(vi) *Representation, familiarity and simplicity considered together.* To some extent the characteristics of representation, familarity, and simplicity are aspects of general ' low information content.' Indeed, half the items in the exhibition had all three characteristics. It is relevant, therefore, to consider the proportion of subjects in each age-group choosing one of these pictures of ' low information content ': the proportions are, in order of increasing age-level, 68 per cent., 71 per cent., 54 per cent., and 39 per cent. Following Maxwell's method (1961), the partitioned χ^2 based on 1 d.f. due to linear regression equals 5·4 $p < \cdot 05$. There is thus an increasing preference for greater ' information content ' with age, but note that 7 and 11-year-olds do not differ in the expected direction.

It is also relevant to consider the relative contribution of each characteristic in determining a picture's popularity with younger children. For this purpose, the seventy-six items are categorised in a 2^3 factorial array, in eight groups, according to their representation, familiarity, and simplicity. The average popularity of each group of items is measured as the proportion, p, liked best by at least one 7 or 11-year-old. This method ensures that effects are not due to any particular picture. The eight p values may be subjected to analysis of variance for a factorial design, using Yates's method. The mean popularity, p, over the eight groups is 33 per cent., and the estimated separate effects are: representation, +20 per cent.; familiarity, +8 per cent.; simplicity, +1 per cent. Representation has thus about twice as much effect as familiarity of content, whilst simplicity has little effect independent of the other characteristics.

(vii) *Brightly coloured versus dull pictures.* All age-classes prefer brightly coloured pictures, and there is little evidence for a decline in this tendency with age (see Table 8). However, older children make about as many ' approving comments ' concerning dull pictures as about brightly coloured ones, and they differ significantly from the 7-year-olds in this respect ($\chi^2=8\cdot 3$ 1 d.f., $p < \cdot 01$). Otherwise, these results confirm those concerning ' reasons for choice,' in showing no decline in the influence of colour with age.

Hildreth (1936) found that even 3–6-year-olds much prefer coloured pictures. Newcomb (reported in Valentine, 1962) found that in writing of the most beautiful thing they had ever seen, 40 per cent. of 7-year-olds mentioned colour, and this proportion even increased with age to 71 per cent. of 13-year-olds. Yet Eysenck (1947) wrote " results seem to indicate a decrease in colour reaction with age." Most results which appear to indicate this do so because the test is such that reaction to form excludes the possibility of reacting also to colour (e.g., Calkins, 1900); and even then results are complicated. Thus, one of the experiments cited by Eysenck (Hazlitt, 1930) actually showed that colour differences in concept-formation cards mainly affected *older* children. Brian and Goodenough (1929) found that young children, given an object

TABLE 8

RELATIVE APPRECIATION OF BRIGHTLY AND DULL-COLOURED PICTURES.

Measure of Appreciation	Subjects	Bright/dull Colour Preference Ratio	Reliability of deviation from Equal Preference	
			χ^2 (1 d.f.)	P<
Individual Preferences	All Children	2·1	9·9†	·01
	Adults	3·9		·05
Approving Comments.	7-year-olds	2·7	9·6	·01
	11 and 15-year-olds	0·9	—	—
Interests Shown	All children	2·0	8·8	·01

† Binomial test used.

matching task, tended to choose the same object as before, showing that they recognised both form and colour. When such perseverative choices were excluded, it was found that 3–6-year-olds tended to match on the basis of colour, whilst under 3-year-olds and adults tended to match on the basis of form !

(viii) *Sex differences.* This analysis applies only to the few items classified as likely to appeal in subject matter especially to boys or to girls. For these items, preferences were noted for boys and girls separately. The association between the subjects' sex and the nature of the pictures was measured by Yule's coefficient : for 7 and 11-year-olds, $A=+0·9$, and for 15-year-olds, $A=+0·6$. The sex difference is significant (for all children combined, $\chi^2=10·3$, 1 d.f., $p<·01$). The particular interests of girls are also shown by the relevant preference-ratio for ' interests shown ' of 3·2 ($\chi^2=6·4$, 1 d.f., $p<·05$). These results conform with previous reports. Munro, *et al.* (reported in Valentine, 1962) found differences in preferences between young boys and girls, which were much less pronounced with older, artistic children. Berry (1942) noted that boys like pictures of ships, whilst girls like fairies and angels.

(*f*) *Prestige suggestion and prferences.*

The results of the prestige suggestion test are unexpected. 77 per cent. of children tested agree with the interviewer's stated preference ($N=13$, Sign Test, $p<·05$) ; whereas 71 per cent. *disagree* with their teacher's suggested preference ($N=21$, Sign Test, 2-tail, $p<·05$). The results for the two independent groups are significantly different (Fisher's exact test, 2-tail, $p<·05$) ; the difference is in the same direction for each age-group. Thus, the stated opinion of the interviewer face-to-face had a positive influence, whilst that of their teacher stated in her absence, tended to give a negative suggestibility effect.

Hovland, *et al.* (1953) noted that face-to-face communication with someone a little older than the listener may be treated as especially credible. Although no age trend is apparent in the present results, previous work suggests a marked decline in prestige suggestibility from about 8 years of age upward (e.g., Eysenck 1947). A few studies report negative suggestibility to authority, particularly in adoelscence (Hull, 1933 ; McConnell, 1963).

ACKNOWLEDGEMENTS.—We are very grateful to the following :

Those who helped in organising the exhibition :

F. W. Hawcroft, Keeper of the Whitworth Art Gallery, and his staff.
T. C. Keay, Organiser of Art for the Manchester Education Committee, who selected the pictures from the Committee's loan collection.
Professor J. White, Pilkington Professor of History of Art.

Interviewers :

Jeanne Callister, P. W. Critten, Suzanne Fowley, M. S. Goodstadt, C. Lillie, G. R. Roberts, Jane Thompson, Helen Williams.

Judges who helped to categorise paintings :

Bridget Peel, Julia Radcliffe.

Judges who helped to categorise responses :

M. S. Goodstadt, Bridget Peel.

Also the teachers and pupils of Heap County Junior, Greenfield County Primary, Ducie Technical High, and Didsbury Technical High Schools, who co-operated so fully in the investigation.

IV.—REFERENCES.

ARNHEIM, R. (1956). *Art and Visual Perception.* London : Faber and Faber.

BARNHART, E. N. (1942). Developmental stages in compositional construction in children's drawings. *J. Exp. Educ.*, 11, 156-184.

BERLYNE, D. E. (1958). The influence of complexity and novelty in visual figures on orienting responses. *J Exp. Psychol.*, 55, 289-296.

BERRY, A. (1942). *Art for Children* (reprint). London : Studio Publications.

BRIAN, C. R., and GOODENOUGH, F. L. (1929). The relative potency of colour and form perception at different ages. *J. Exp.* Psychol., 12, 197-213.

BULLEY, M (1934). Aesthetic judgments of children. *Brit. J. Educ. Psychol.*, 4, 162-185.

CALKINS, M. (1900). An attempted experiment in psychological aesthetics. *Psychol. Rev.*, 7, 580-591.

DASHIEL, J. F. (1917). Children's sense of harmonics in colours and tones. *J. Exp. Psychol.*, 2, 466-475.

DEWAR, H. (1938). A comparison of tests of artistic appreciation. *Brit. J. Educ. Psychol.*, 8, 29-49.

EYSENCK, H. J. (1940). The general factor in aesthetic judgments. *Brit. J. Psychol.*, 31, 94-102.

EYSENCK, H. J. (1941a). Type factors in aesthetic judgments. *Brit. J. Psychol.*, 31, 262-270.

EYSENCK, H. J. (1941b). The empirical determination of aesthetic formula. *Psychol. Rev.*, 48, 83-92.

EYSENCK, H. J. (1942). The experimental study of the ' good gestalt '— a new approach. *Psychol. Rev.*, 49, 344-364.

EYSENCK, H. J. (1947) *Dimensions of Personality.* London · Kegan Paul.

FRENCH, J. E. (1952). Children's perferences for pictures of varied complexity of pictorial pattern. *Elem. Sch. J.*, 53, 90-95.

GORDON, D. A. (1955). Artistic excellence of oil paintings as judged by experts and laymen. *J. Educ. Res.*, 48.

HARRIS, D. B. (1963). *Children's Drawings as Measures of Intellectual Maturity.* London : Harrap.

HAZLITT, V. (1930). Children's thinking. *Brit. J. Psychol.*, 20, 354-361.

HILDRETH, GERTRUDE (1936). Colour and picture choices of young children. *J. Genet. Psychol.*, 49, 427-435.

HILDRETH, GERTRUDE (1944). The simplification tendency in reproducing designs. *J. Genet. Psychol.*, 64, 329-333.

HOVLAND, C. I., JANIS, I. L., and KELLEY, H. H. (1953). *Communication and Persuasion.* New Haven : Yale Univ. Press.

HULL, C. L. (1933). *Hypnosis and Suggestibility.* New York : Appleton-Century.

LIPSITT, L. P., and SPIKER, C. C. (Eds.) (1963). *Advances in Child Development and Behaviour.* Vol. 1. London : Academic Press.

LITTLEJOHNS, J., and NEEDHAM, A. (1933). *Training of Taste in the Arts and Crafts.* London : Pitman.

McCONNELL, T. R. (1963). Suggestibility in children as a function of chronological age. *J. Abnorm. Soc. Psychol.*, 67, 286 289.

McGLYNN, T., and DEARNALEY, E. J. (1963). Some changes in the paintings of students attending discussion and criticism classes. ***Durham Res. Rev.***, 14, 77-82.

MAXWELL, A. E. (1961). ***Analysing Qualitative Data.*** London : Methuen.

MEIER, N. C., *et al.* (1933). Studies in the psychology of art, I. *Psychol. Monogr.*, 45, Whole No. 200.

PEEL, E. A. (1945). On identifying aesthetic types. *Brit. J. Psychol.*, 35, 61-69.

PICKFORD R. W. (1948). 'Aesthetic' and 'technical' factors in artistic appreciation. *Brit. J. Psychol.*, 38, 135-141.

SPITZ, H. H., and HOATS, D. L. (1961). ***Experiments on Perceptual Curiosity Behaviour in Mental Retardates.*** Final report, N.I.M.H. Grant M/4533, Johnstone Training and Research Centre.

TOMLINSON, R. R. (1947). ***Children as Artists*** 2nd ed. London : Penguin Books.

VALENTINE, C. W. (1962). *The Experimental Psychology of Beauty.* London : Methuen.

VERNON, M. D. (1940). The relation of cognition and phantasy in children, I. *Brit. J. Psychol.*, 30, 273-294.

WINSTANLEY, B. R. (1963). ***The Encouragement of Appreciation in Children.*** Paper read at the "Pictures for Schools" Course, Manchester.

(***Manuscript received*** 24***th March***, 1966)

DEVELOPMENTAL CHANGES IN CHILDREN'S DRAWINGS

BY MARTIN P. M. RICHARDS AND HELEN E. ROSS*
(University of Cambridge)

SUMMARY. A drawing of a cat and kittens under standard conditions was collected from about 1,200 normal school children aged 5–14 years. Thirty boys and thirty girls of similar intelligence were selected from each year of age. Their drawings were then scored on various objective measures, such as the area filled and the number of colours used. Significant differences were found between the two sexes, among younger children: the girls tended to produce more 'mature' drawings than the boys. Complex age trends were found, several measures showing a regression to a more childish style of drawing after the age of 11–12 years.

I.—INTRODUCTION.

THE drawings and paintings of normal and abnormal children and adults have been used for assessing personality, measuring intelligence and predicting mental disorder. The large literature in this field has been extensively reviewed (Goodenough, 1926; Eng., 1931; Anastasi and Foley, 1941; Goodenough and Harris, 1950; Naville, 1950; Eng., 1957; Johnson and Gloye, 1958). Much of the published work has been criticised because of subjective measurements, and inadequate control groups and statistical treatment. This study was designed to investigate sex differences and developmental changes in the drawings of normal children, produced under standard conditions and scored objectively.

II.—METHOD.

Subjects.

Drawings were collected from about 1,200 children attending five local authority schools. These were a small primary and infant school in the centre of Cambridge; a large primary and infant school on a housing estate in Cambridge; a boys' and a girls' secondary school on the same housing estate; and a village college in Cambridgshiree. The children's ages ranted from 4 to 11 years in the infant and primary schools, and from 11 to 15 in the secondary schools and village college. Apart from the two secondary schools, all classes contained children of both sexes.

Testing Procedure.

As far as possible, testing was carried out in the normal school art periods Each child was provided with a sheet of grey sugar paper 12½-in. × 20-in., and a box containing nine Reeves' pastels. The colours were black, white, blue, green, yellow, orange, red, purple and brown. The children were then asked to draw a cat and kittens, and told that they had 25 minutes in which to complete the picture. This subject was chosen for the drawings because most children are familiar with it, and because it had given consistent results in a pilot study. No other instructions or help were given. The name of each child was recorded, and also the age (in years and months), sex, I.Q. (if known from any previous tests), and any unusual medical or psychiatric history. The class teacher was also asked to rate the child's intelligence as high, average or low.

---

* Now at Department of Psychology, University of Hull.

Selection of Sample.

For each year of age from 5 to 14, a group of thirty boys and thirty girls was selected. Each group of thirty children was composed of three sub-groups of ten. Each sub-group was selected from children of 'high,' 'average' or 'low' intelligence. The grading of the children was based on the available information from previous intelligence tests and the teachers' estimates. Children were used in the sample only if there was general agreement between the two sources of information.

No children were included in the sample who suffered from serious visual defects or colour blindness, or who were considered abnormal for medical or psychiatric reasons. The groups were composed of children from as many schools and classes as possible, to minimize the effects of differences in art instruction. All groups contained children from at least two schools and three classes.

The mean age of each group was always between five and seven months beyond the year quoted (so that age 11 means about 11½ years).

The I.Q. information was not used in the final analysis, since it was based upon a variety of different tests which were difficult to compare. A preliminary analysis suggested that some differences in drawing-scores were correlated with differences in intelligence; but these differences were generally very small in comparison with those due to age and sex. The I.Q. groups were therefore combined, and the present analysis is concerned only with age and sex as variables.

Measurement of Drawings.

The drawings were scored under the following headings:

(*a*) *Colours used.* The number of colours used in each drawing was recorded regardless of the amount of use of each colour. The actual colours used were also recorded.

(*b*) *Unrealistic colour.* This referred only to the cat and kittens, and not to any other objects included in the picture. The following colours were counted as realistic: black, white and brown anywhere on the cat; orange and yellow on the body; blue, green and yellow on the eyes; red on the ears and lips. If other colours were used, the picture was counted as containing unrealistic colour.

(*c*) *Area of drawing.* Each drawing was placed under a transparent plastic grid divided into sixty squares. The number of squares containing any crayon marks was counted. A separate count was made of the number of squares on the periphery of the paper containing crayon marks.

(*d*) *Background.* Many children drew some background or extra objects in addition to the cat and kittens. The number of pictures with such background was recorded.

(*e*) *Strips of sky and ground.* Many young children drew a strip of blue at the top of the paper indicating the sky, and a strip of green or brown at the bottom of the paper indicating the ground. Sometimes bizarre colours were used for the strips. The pictures were scored with one point if either sky or ground was indicated by a strip, and two points if both were indicated in this way.

(*f*) *Outlining.* 'Pure outlining' was defined as leaving the body of the cat empty of colour. 'Differential outlining' was defined as filling in the body of the cat in a different colour from the outline, except where the outline colour was black, white or brown (children of all ages tended to treat these colours as pencil outlines). The number of pictures containing pure outlining and differential outlining was recorded separately.

III.—Results.

In most of the measurements used in this study there was an upper limit to the scores obtainable, or the scores were in categories which do not form a normal distribution. For this reason, parametric statistics were inappropriate, and the median or chi-squared test was used to test the differences between the sex groups. The significance of the age-trends is difficult to evaluate on this basis ; but the fact that the same age-trends were present for both sex groups suggests that they were not due to chance.

(*a*) *Colours used.* The mean number of colours used per drawing is shown in Table 1 for each age/sex group. At all ages up to 11, the girls used more colours than the boys, but there were no significant differences among the older children. The difference between the sexes was tested on the median test, dividing the children into groups aged 5–7, 8–11 and 12–14. This gave χ^2 values of 20·53 ($p < 001$) and 4·68 ($p < \cdot 05$) for the first two troups, and 1·098 (not significant) for the oldest children. It can be seen from Table, 1 that the number of colours used rose to a peak at age 7 or 8, declined slightly, reached a further peak at 11, and then declined sharply. As the two-peaked change with age occurred for both sexes, it is unlikely to be due to random fluctuations. There is some suggestion that the girls reached their peaks earlier than the boys.

TABLE 1

Mean Number of Colours Used Out of a Maximum of 9.

Age	Boys	Girls
5	5·4	6·7
6	5·3	6·9
7	6·1	7·5
8	6·9	7·2
9	6·0	6·8
10	6·3	7·6
11	7·3	7·8
12	6·7	5·9
13	5·5	5·9
14	5·7	4·9

A measure of group colour preferences was obtained by calculating the number of times each colour was used as a percentage of the total number of colours used by each group. The results are shown in Table 2. As nine colours were available, colours selected randomly scored 11·1 per cent., while colours preferred or ignored scored above or below that percentage. The colours fell into four groups : those which were selected positively (black, white, brown) ; those treated neutrally (blue and green) ; those treated neutrally or slightly negatively (yellow and red) ; and those treated almost entirely negatively (purple and orange). At all ages, except 12 and 14, the boys showed more selection than the girls. This is indicated by the greater difference of their scores from the random level. The ratio of the most to the least preferred colour gives a general measure of selectivity, and is shown in the last column of Table 2. Both sexes showed the same age trends : at age 5 selection was almost random ; it then increased to a peak at 9, decreased to a more random level at age 11, then increased sharply to age 14. There was again a suggestion that the girls' peaks and dips occurred slightly earlier than the boys'.

TABLE 2

PREFERENCES FOR DIFFERENT COLOURS.

The total number of children using each colour is expressed as a percentage of the total for all the colours, for each age group. The colours are arranged in approximate order of preference from left to right. Colours scoring 11·1 per cent. were selected randomly, while those scoring above or below that percentage were preferred or rejected. The last column shows the ratio of the most to the least preferred colour, as a general measure of selectivity.

Age	Sex	Black	White	Brown	Blue	Green	Yellow	Red	Orange	Purple	Preference Ratio
5	B	10·5	14·2	8·4	11·1	13·0	11·7	10·5	9·9	10·5	1·65
	G	10·9	13·9	9·9	11·4	10·9	10·9	10·9	8·9	12·4	1·56
6	B	14·6	15·2	14·6	10·1	10·8	10·1	8·2	8·2	8·2	1·85
	G	11·5	13·5	10·1	11·5	11·0	12·5	12·0	9·6	8·2	1·65
7	B	14·7	15·8	14·1	12·5	9·8	11·4	8·7	7·6	5·4	2·93
	G	12·9	12·9	12·5	12·1	9·4	11·2	12·5	7·6	8·9	1·70
8	B	14·6	14·6	14·1	10·7	11·7	10·2	9·7	8·3	6·3	2·30
	G	12·4	13·8	12·9	10·6	10·6	11·5	9·2	11·1	7·8	1·77
9	B	15·6	16·8	14·0	11·2	12·3	8·4	9·5	8·4	3·9	4·28
	G	14·3	14·3	13·3	11·3	12·8	8·7	7·9	10·3	6·9	2·07
10	B	15·9	15·9	14·2	8·9	8·9	10·0	11·1	10·5	4·7	3·34
	G	13·2	13·2	11·5	10·1	10·8	10·8	11·9	8·8	10·1	1·5
11	B	13·1	12·7	10·9	9·5	12·7	12·3	10·0	9·5	9·1	1·46
	G	11·9	12·8	11·5	10·2	12·3	10·2	10·2	9·4	11·5	1·36
12	B	14·4	13·9	12·4	10·4	10·9	11·4	10·9	7·9	7·9	1·81
	G	15·7	15·7	14·0	11·2	12·4	11·2	7·9	6·7	5·1	3·08
13	B	16·4	17·0	14·5	9·1	10·9	10·9	9·7	7·3	4·2	4·02
	G	15·7	16·3	14·6	8·4	13·5	9·6	7·9	9·6	4·5	3·62
14	B	17·0	17·5	15·2	8·2	9·9	11·7	8·8	7·6	4·1	4·38
	G	19·7	18·4	16·3	12·2	11·6	4·8	3·4	11·6	2·0	9·85

TABLE 3

PERCENTAGE OF CHILDREN IN EACH AGE-GROUP USING UNREALISTIC COLOUR.

Age	Boys	Girls
5	93·3	93·3
6	63·3	80·0
7	36·7	50·0
8	40·0	23·3
9	30·0	6·7
10	43·3	10·0
11	30·0	3·3
12	23·3	0·0
13	6·7	6·7
14	20·0	6·7

(*b*) *Unrealistic colour.* The percentage of boys and girls using unrealistic colour is shown in Table 3. For the age-group 5–7, the girls used slightly more unrealistic colour than the boys, but this was not significant ($\chi^2=1{\cdot}68$). For the age-groups 8–11 and 12–14, the boys used significantly more unrealistic colour than the girls ($\chi^2=5{\cdot}24$ $p<{\cdot}05$; $\chi^2=6{\cdot}80$ $p<{\cdot}02$). Its use declined with age among both sexes ; but there was a suggestion of an increase among the girls after age 12, and among the boys after age 13.

(*c*) *Area of drawing.* The same trends were shown whether the total number of filled squares was counted, or only those on the periphery. As the latter measure showed the trends more clearly, it is used here. Table 4 shows the differences between the sex groups in the percentage of the periphery filled. Girls filled up more of the outer edges of the paper than the boys up to the age of 10, after which the trend reversed. If these differences are tested in three age groups as before, they fail to be significant. However, if they are divided into two groups (5–9 and 10–14), the younger girls filled in significantly more squares than the younger boys ($\chi^2=4{\cdot}34$ $p<{\cdot}05$). There were no significant differences in the older group. The age trends were similar to those found with colours : there was a peak at 7 or 8, and at 11 or 12.

TABLE 4

PERCENTAGE OF PERIPHERAL AREA OF PAPER FILLED BY EACH AGE-GROUP.

Age	Boys	Girls
5	28·8	68·3
6	36·3	54·4
7	63·4	64·6
8	61·6	63·8
9	46·4	50·1
10	47·4	49·9
11	81·0	63·0
12	74·3	75·5
13	57·1	51·7
14	65·8	44·3

(*d*) *Background.* Table 5 shows the percentage of boys and girls drawing cats with backgrounds. Until the age of 11, the boys drew fewer backgrounds

TABLE 5

PERCENTAGE OF CHILDREN IN EACH AGE-GROUP ADDING BACKGROUND.

Age	Boys	Girls
5	46·7	66·6
6	56·7	80·0
7	86·7	90·0
8	93·3	96·7
9	76·7	96·7
10	90·0	96·7
11	100·0	100·0
12	92·7	93·4
13	83·3	83·4
14	93·3	76·7

than the girls, after which there was no clear trend. The difference was significant for the 5–7 age-group only ($\chi^2=4{\cdot}57$ $p<{\cdot}05$). There was a steady increase in the number of backgrounds until the age of 9 and a decline from age 12.

(*e*) *Strips of sky and ground.* Table 6 shows the total number of strips drawn by the boys and girls. The girls drew more strips than the boys at all ages ($\chi^2=14{\cdot}98$ $p<{\cdot}001$). The drawings of strips ceased entirely after age 9, and reached a peak among the girls at age 6 and the boys at age 7.

TABLE 6

TOTAL NUMBER OF STRIPS OF SKY OR GROUND DRAWN BY EACH AGE-GROUP.

Age	Boys	Girls
5	6	17
6	17	28
7	18	19
8	5	20
9	2	6
10	0	0
11	0	0
12	0	0
13	0	0
14	0	0

(*f*) *Outlining.* The scores were noted both for 'pure outlining' and for 'differential outlining.' These are shown for both sexes in Table 7. The sexes were very similar on differential outlining. The boys tended to use more pure outlining than the girls until the age of 12, when the trend reversed. The differences were not significant, but very small numbers were involved. Pure outlining was entirely absent among the girls from the ages 7-10, inclusive, but reappeared from 11 on. Differential outlining declined sharply to age 9, but never disappeared completely. Up to age 10 differential outlining always exceeded pure outlining, but at later ages their incidence was similar.

TABLE 7

NUMBER OF CHILDREN IN EACH AGE-GROUP USING PURE AND DIFFERENTIAL OUTLINING.

Age	Boys		Girls	
	Pure	Differential	Pure	Differential
5	13	15	3	23
6	3	12	2	12
7	2	9	0	9
8	0	3	0	4
9	1	2	0	2
10	0	1	0	1
11	2	2	1	0
12	1	1	1	0
13	1	1	4	1
14	1	2	3	2

IV.—DISCUSSION.

All the measures used showed changes with age and sex. It is, perhaps, surprising that no measures showed a simple age-trend with no reversals. Most measures showed a reversal of trend around the age of 12 (except for the drawing of strips, which had ceased entirely by that age). Some measures reached a peak at the age of about 12, and declined thereafter: these were the number of colours used, the area filled, and the drawing of backgrounds. Others reached a minimum and then increased: these were the use of unrealistic colour (at 12–13), and pure outlining (at 8–10).

Several authors (e.g., Burt, 1921; Tomlinson, 1944) have noted that the drawings of adolescents tend to deteriorate or resemble those of younger children. The ages at which this deterioration, and other developmental changes, are said to occur are reviewed by Weil (1950). It is often suggested that this regression is due to a realistic and critical attitude which arrives with puberty and depresses creative ability. However, we have no data which would enable us to test the theory that the regression is caused by the onset of puberty. It is possible that some deterioration occurs as a function of age regardless of the onset of puberty.

In addition to the peak in drawing ability at about 11–12, some measures show an earlier peak at about 7–8 followed by a trough at 9–10. This trend is evident in the number of colours used, the area filled, and (perhaps) in the addition of background. Several authors have commented on the halt, or dip, in drawing ability at about 8–10. Peck (1936) claimed that the increase in ability to copy photographs slowed up after the age of 8, and Cameron (1938) and Löwenfeld (1939), both mentioned deterioration at this age. Childs (1915) found peaks in drawing ability at 8 and 11–12. Löwenfeld attempted to explain the dip by saying that at this age the child becomes aware of his environment and loses faith in his creative abilities; while Cameron stated that this was the age when the child's interest shifted from visual and verbal to social activities. These theories are hard to test, and they do not explain why some aspects of drawing ability increase steadily with age (e.g., Goodenough, 1926).

The measure of group colour selectivity shows complex trends. As might be expected, it shows some inverse correlation with the number of colours used. Selectivity reaches a peak at about 9 and 14, when the number of colours used shows a dip; and it reaches a minimum at 11, when the number of colours used is approaching a peak. However, the correlation is not complete: at 5 few colours are used, nearly randomly; while at 14 few colours are used, highly selectively. It would be desirable to find some measure of colour selectivity within each drawing rather than within a group of drawings. If the group measure reflects the selectivity of individuals, it shows that at least one aspect of artistic ability increases steadily after the age of 11. Alternatively, it may only indicate that adolescents are more conformist than 5-year-olds about which colours to use.

The measures concerned with the use of colour show the clearest sex differences. Girls use more colours than boys, but with less selectivity (or less conformity). The change with age in both these measures tends to occur about a year earlier in girls than boys; perhaps this reflects the earlier maturity of girls. Other authors (e.g., Burt, 1921; Altschuler and Hattwick, 1947) have reported that girls use more colour than boys. Eng (1957), in her study of her niece's drawings and paintings, found the number of colours used increased to age 7, then dropped and rose again from 10 to 11, after which form apparently became more important than colour. This agrees well with the changes we have found. The order of colour preferences we have obtained does not agree with Eng's results, nor with the order found by Eysenck (1941) from preference tests. However, the order obtained is likely to vary considerably with the subject matter of the picture or the nature of the preference test.

Young girls tend to fill up more of the paper and add more background than boys. This largely took the form of the addition of strips of sky and ground, which ceased after the age of 9 when the sex differences disappeared in the area filled and the addition of background. It is not clear why girls should produce more of these pictures with stereotyped sky and ground.

ACKNOWLEDGEMENTS.—We should like to thank the following people for their advice and co-operation : Mr. D. A. F. Conochie, M.A., Educational Psychologist for the City and County of Cambridgeshire ; Mr. G. D. Edwards, M.A., Chief Education Officer for Cambridgeshire ; Mr. M. M. Scarr, G.M., M.A., City Education Officer for Cambridge ; the head masters, head mistresses, form teachers, art teachers and school doctors of the following schools : Manor Secondary Boys' School, Manor Secondary Girls' School, Impington Village College, Arbury Junior School, Arbury Infants' School, Park Street C. of E. Primary School, Dry Drayton Primary School. We should also like to thank Professor R. A. Hinde for advice and encouragement and Professor P. Bertelson and Mrs. E. Barton for statistical and other help.

This research was carried out while both authors held research studentships from the Department of Scientific and Industrial Research.

V.—REFERENCES.

ALTSCHULER, R. H., and HATTWICK, LA B. W. (1947). *Painting and Personality : A Study of Young Children*, Vols. 1 and 2. Chicago : Univ. of Chicago Press.

ANASTASI, A., and FOLEY, J. P. (1941). A survey of the literature on artistic behaviour, in the abnormal. 1—Historical and theoretical background. *J. Gen. Psychol.* 25, 111-142.

BURT, C. (1921). *Mental and Scholastic Tests.* London : King.

CAMERON, N. (1938). Individual and social factors in the development of graphic symbolism. *J. Psychol.*, 5, 165-184.

CHILDS, H. G. (1915). Measurement of the drawing ability of 2,177 children in Indiana City school systems by a supplemented Thorndike scale. *J. Educ. Psychol.*, 6, 391-408.

ENG, H. (1931). *The Psychology of Children's Drawings.* London . Kegan Paul.

ENG, H. (1957). *The Psychology of Child and Youth Drawing.* London : Routledge and Kegan Paul.

EYSENCK, H. J. (1941). A critical and experimental study of colour preference. *Amer. J. Psychol.*, 54, 386-394.

GOODENOUGH, F. L. (1926). *Measurement of Intelligence by Drawings.* New York : World Book Co.

GOODENOUGH, F. L., and HARRIS, D. B. (1950). Studies in the psychology of children's drawings, II, 1928-1949. *Psychol. Bull.*, 47, 369-343.

LOWENFELD, V. (1939). *The Nature of Creative Activity.* New York : Harcourt, Brace and Co.

NAVILLE, P. (1950). Éléments d'une bibliographie critique relative au graphisme enfantin jusqu'en, 1949. *Enfance*, 3, 310-407.

PECK, L. (1936). An experiment in drawing in relation to the prediction of school success. *J. Appl. Psychol.*, 20, 16-43.

TOMLINSON, R. R. (1944). *Children as Artists.* London . Penguin Books.

WEIL, P. G. (1950). Caractéristiques du développement du dessin per groupes d'ages, selon divers auteurs. *Enfance* 3, 221-226.

(*Manuscript received 24th March*, 1966)

INTELLIGENCE ASSESSMENT: A THEORETICAL AND EXPERIMENTAL APPROACH*

By H. J. EYSENCK,

(*Institute of Psychiatry, University of London*)

I.—DEVELOPMENT OF A CONCEPT.

ATTEMPTS to measure intelligence have passed through several stages since Galton tried to use the measurement of sensory processes to arrive at an estimate of the subject's intellectual level (1883), and McKeen Cattell 1890) employed tests of muscular strength, speed of movement, sensitivity to pain, reaction time and the like for a similar purpose. These largely abortive efforts were followed by the first stage of intelligence measurement properly so called; it may with truth be labelled the 'g' phase because both Spearman (1904) and Binet and Simon (1905) stressed the importance of a *general factor of intellectual ability*, Binet contributing mainly by the construction of test items and the invention of the concept of mental age, Spearman contributing mainly by the application of correlational methods and the invention of factor analysis.

The second stage was concerned with the proper definition of intelligence, and theories regarding its nature. Several books concerned themselves with this problem (Thurstone, 1926; Spearman, 1923), and a number of symposia were held (*Brit. J. Psychol.*, 1910; *J. Educ. Psychol.*, 1921; *Internat. Congress of Psychol.*, 1923). Among the theories canvassed were 'mental speed' hypotheses which placed the burden of intellectual attainment on speed of mental functioning, and 'learning' hypotheses which protested that the ability to learn new material was fundamental. Both hypotheses faced difficulties; the fact that reaction times showed no relation to ability tended to discourage believers in the 'speed' hypothesis, and the negative results of the large-scale work of Woodrow (1946) on the relation between different learning tasks and intelligence discouraged believers in the 'learning' hypothesis. Psychologists learned to agree to disagree, and to present their work with the dictum that "intelligence is what intelligence tests measure"—a saying less circular than it sounds, but only acceptable if all intelligence tests did, in fact, measure the same thing, which they quite emphatically did not.

We thus reach the third stage, which is essentially a continuation of the early factor analytic approach, but now fortified by recourse to multiple factors and matrix algebra. This phase owes most to Thurstone, but Thomson, Burt, Holzinger, and many others made valiant contributions. In this factorial phase, investigators went back to Binet's idea of different mental faculties making up the complex concept of intelligence, and used factor analysis to sort out these alleged faculties; they emerged with verbal, numerical, perceptual, memory, visuo-spatial and many other factors. At first, Thurstone and his followers believed that these 'primary factors' put paid altogether to the notion of intelligence, but when they found the primary factors to be themselves correlated they resurrected the concept of intelligence as a second-order factor, a solution already implicit in the earlier methods and theories of Burt (Eysenck, 1939).

---

* This paper was originally delivered at a symposium on New Aspects of Intelligence Assessment at the Swansea Meeting of the B.P.S., on 3rd April, 1966. The preparation was assisted by a grant from the M.R.C.

The fourth stage constitutes essentially an extension of the third, and is associated specifically with J. P. Guilford (1966), whose publication of his " 1965 model of intelligence " provided some of the motivation for this paper. This model, which shows some similarities to one I published in Uses and Abuses of Psychology (1953, p. 38), is illustrated in Fig. 1. Guilford classifies the intellect into *operations* which it can perform, different *contents* of these operations, and different *products* ; by taking all possible interactions we obtain 120 cells corresponding to different mental abilities. Of these Guilford claims to have evidence in actual factorial studies for eighty ; he is optimistic about discovering the remainder. To some critics, this factorial extension of Thurstone's work has appeared almost as a *reductio ad absurdum* of the whole approach. There is a possibility of infinte sub-division inherent in the statistical method employed, and evidence is lacking that further and further sub-factors add anything either to the experimental analysis of intellectual functioning or the practical aim of forecasting success and failure in intellectual pursuits (Vernon, 1965). Worse, the model fails to reproduce the essentially hierachical nature of the data ; the one outstanding fact which recurs again and again in all analyses is the universality of positive correlations among all relevant tests, and the positive correlations between different factors (McNemar, 1964). By omitting any mention of this central feature of the scene Guilford has truly cut out the Dane from his production of Hamlet. If this is really the best model (1965 style) which psychology can offer of intelligence and intellect, then the time seems to have come to retrace our steps ; something has gone very wrong indeed !

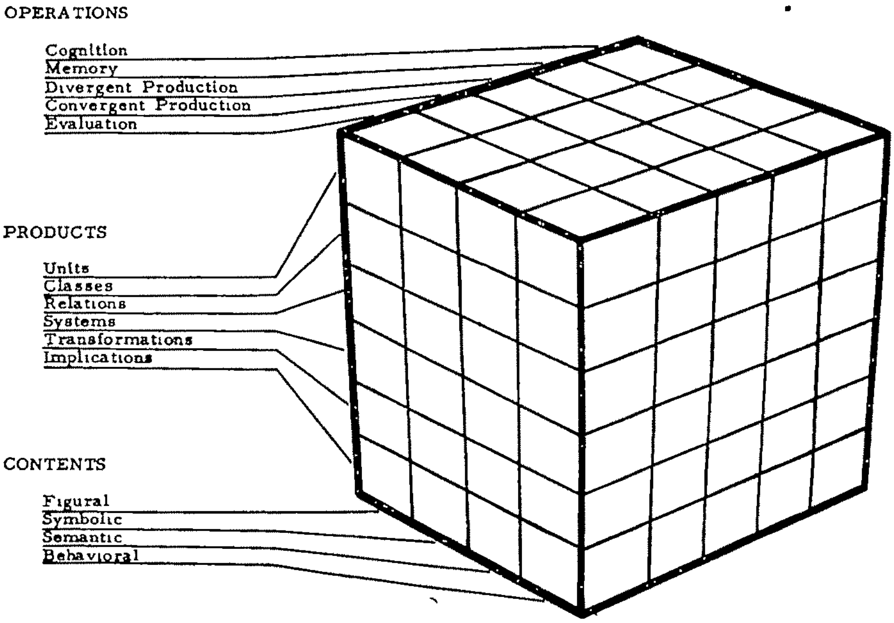


FIGURE 1.

Model of the structure of intellect (Guilford, 1966).

II.—LIMITATION OF THE FACTOR ANALYSIS APPROACH.

Zangwill has several times suggested that the whole intelligence testing movement is a technological rather than a scientific one, and in essence my own diagnosis is not too different from his. I would suggest that the psychometric approach has become almost completely divorced from both psychological theory and experiment, and that factor analysis, while an extremely useful tool, cannot by itself bear the whole burden which has been placed upon it. It is the purpose of this paper to raise certain questions in this connection rather than to give definitive answers ; a few empirical results from some of our work will be presented more in order to illustrate an approach than because we believe that these results settle the questions the experiments were designed to investigate.

TABLE 1

FIVE-ITEM INTELLIGENCE TEST, ADMINISTERED TO FIVE CHILDREN ALL HAVING A SCORE OF 2.

	1	2	3	4	5	Total Score
Jones	R	R	N	N	N	2
Charles	W	R	W	R	N	2
Smith	R	A	A	R	A	2
Lucy	R	A	N	N	N	2
Mary	R	W	R	W	W	2

R—Right answer. W—Wrong answer. A—Abandoned item. N—Item not attempted. (In most tests A and N cannot be distinguished.)

Our work started out with a fundamental criticism of the whole testing movement, directed at the unit of analysis chosen. Nearly all factor analysts and psychometrists correlate test scores and then proceed to work with these correlations ; they thus assume that equal scores are equivalent. Such an assumption is unwarranted in the absence of proof, and consideration of typical intelligence test papers shows that it is, in fact, mistaken. Consider Table 1, which shows the results of giving an imaginary five-item test to five candidates. Let R stand for an item correctly solved, W for an item incorrectly solved, A for an item abandoned, and N for an item not attempted. Let us also assume that the items increase in difficulty. It will be seen that all five children obtain an identical mark of 2 ; but it will also be seen that no two children obtain this mark in the same way. Jones gets the easiest two right, but uses up all his time and does not attempt any more ; he works slowly and carefully. Charles gets some easy items wrong and some difficult ones right ; he works quickly but carelessly. Smith gives up on three items ; had he been more persistent, he might have solved some of them. Lucy is rather selective in the choice of item to be tackled, and Mary fails to check her answers, getting three of them wrong. Can it really be maintained that the mental processes and abilities of these five children are identical, merely because they all obtained the same final mark ? This is the implicit assumption underlying the factor analysis of test scores, and it may be suggested that this assumption requires careful investigation before we can regard it as acceptable. Such investigations are notable by their absence, and factor analysts proceed throughout as if the problem did not exist. This, it may be suggested, is not a proper scientific procedure.

III.—The furneaux model.

Our own approach has been to emphasize the point that the fundamental unit of analysis must be the individual test item, and that in addition to determining the category (R, W, A, N) into which it falls for each candidate, it is important to determine the *speed* with which each R item is solved, the length of time devoted to each A item (persistence or continuance), and the number of W items together with the time spent on each. Furneaux (1960) has given a detailed analysis of scores obtained in this fashion, and has suggested on the basis of this evidence that the solution of mental test problems has three main parameters : (1) mental speed, i.e., speed of solution of R items ; (2) Continuance, or persistence in efforts to solve problems the solution to which is not immediately apparent ; and (3) Error Checking Mechanism, i.e., a mental set predisposing the individual to check his solution against the problem instead of writing it down immediately. Two interesting and important consequences follow from this analysis. In the first instance, Furneaux reinstates the mental speed factor to its theoretical pre-eminence as the main cognitive determinant of mental test solving ability, and in the second instance he emphasizes the importance of non-cognitive (personality) factors in determining mental test performance—both persistence and carefulness in checking are personality attributes rather than cognitive abilites. I have attempted to incorporate some elements of this analysis into my own model of intellect (Eysenck, 1953), which is shown in Fig. 2, and which may be compared with Guilford's. What I call ' mental processes ' he calls ' operations ' ; what I call ' test material ' he calls ' contents ' ; so far there is close agreement. But instead of having a third dimension concerned with ' products' (which seems to me a weak and not very important principle of division) I have suggested a dimension rather vaguely labelled ' quality ' into which I wanted to incorporate concepts of mental speed and power, somewhat after the fashion of Thorndike's fundamental contribution (1926). The suggestion is that mental speed and power are fundamental aspects of all mental work, but that they are to some extent qualified by the mental processes involved and the materials used. This seems to me a more realistic concept than Guilford's, as well as having the advantage of retaining the central ' g ' concept in a hierarchical structure in which the

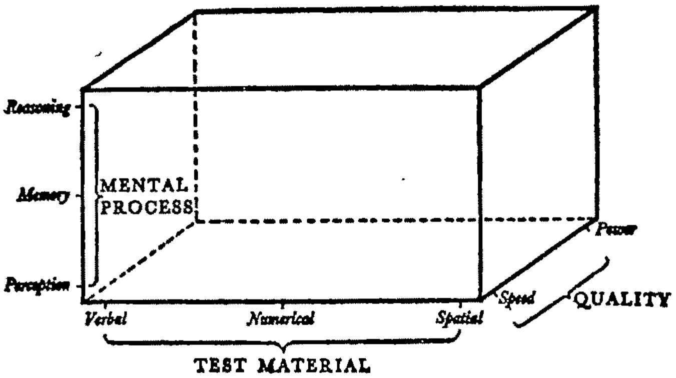


Fig. 2

Figure 2.
Model of the structure of intellect (Eysenck, 1953).

major source of variation is mental speed, averaged over all processes and materials. ' Primary mental abilities,' so called, would then emerge at a lower level of generality, and be related to different processes and different materials used.

Furneaux has demonstrated the fundamental nature of the mental speed function by showing that when an individual's R latencies are plotted against the difficulty level of the items concerned, a negatively accelerated curve is obtained (Fig. 3, A) ; when the time units are then logarithmically transformed *all plots become linear and parallel* (Fig. 3, B). This may be interpreted to mean that the only source of difference in intellectual ability between individuals (in relation to the particular set of test items chosen at least) is the intercept on the abscissa. The increase in log. latency with increase in item difficulty turns out to have the same slope for all individuals tested, and is thus a *constant*, one of the few which exist in psychology. It seems to me that the scientific study of intelligence would gain much by following up the important leads given by Furneaux in this extremely original and path-breaking work.

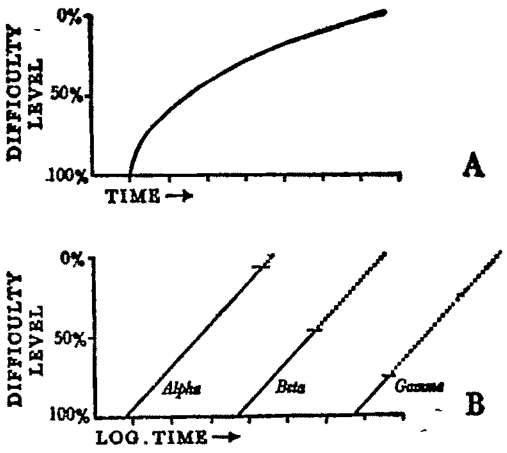


FIGURE 3.

Relation between difficulty level of test items and time (A) and log time (B) needed for solution. Alpha, beta and gamma are three imaginary subjects of high, medium and low mental ability, respectively (Eysenck, 1953).

IV.—MENTAL SPEED AND INTELLIGENCE.

On the theoretical side Furneaux has suggested that what may be involved in problem solving activity may be some kind of scanning mechanism the speed of which determines the probability of the right solution being brought into focus more or less quickly. If we join this notion with that of information processing, we may have here not only the suggestion of a useful theory of intellectual functioning, but also an argument against those who abandoned the whole theory of ' speed ' as underlying intelligence because of the failure of reaction time experiments to correlate with intelligence tests. Let us consider the amount of information conveyed by flashing a light and requiring the subject to press a button located underneath the light flashed. When there is only one light/button combination, no information is, in fact, conveyed. As the number of combinations increases, the amount of information conveyed increases logarithmically, so that one bit of information is conveyed with two combinations, two bits with four combinations and 3 bits with eight combinations. Response speed has been shown by Hick (1952), Hyman (1953) and Schmidtke (1961), to increase linearly with increasing number of bits of information, as shown in Fig. 4 (Frank, 1963). We have two separate items of information for each subject: one is the raw reaction time, as shown by the intercept on the ordinate, the other is the slope of the regression line, i.e., the rate of increase in reaction time with increasing amount of information processed. If intelligence is conceived of as speed of information processing, then simple reaction time, involving 0 bits of information, should not correlate with intelligence, but the slope of the regression line, showing increase of reaction time with amount of information processed, should correlate (negatively) with intelligence; in other words, intelligent subjects would show less increase in reaction time with increase in number of light/button combinations than would dull ones (This is a slightly more precise way of phrasing Spearman's first noegenetic law.) Experimentally, the prediction has been tested by Roth (1964) who demonstrated that while as expected simple reaction time was independent of I.Q., speed of information processing (slope) correlated significantly with I.Q., in the predicted direction. Reaction time experiments, properly interpreted, do not appear to contradict a theory of intellectual functioning based on the motion of mental *speed*.

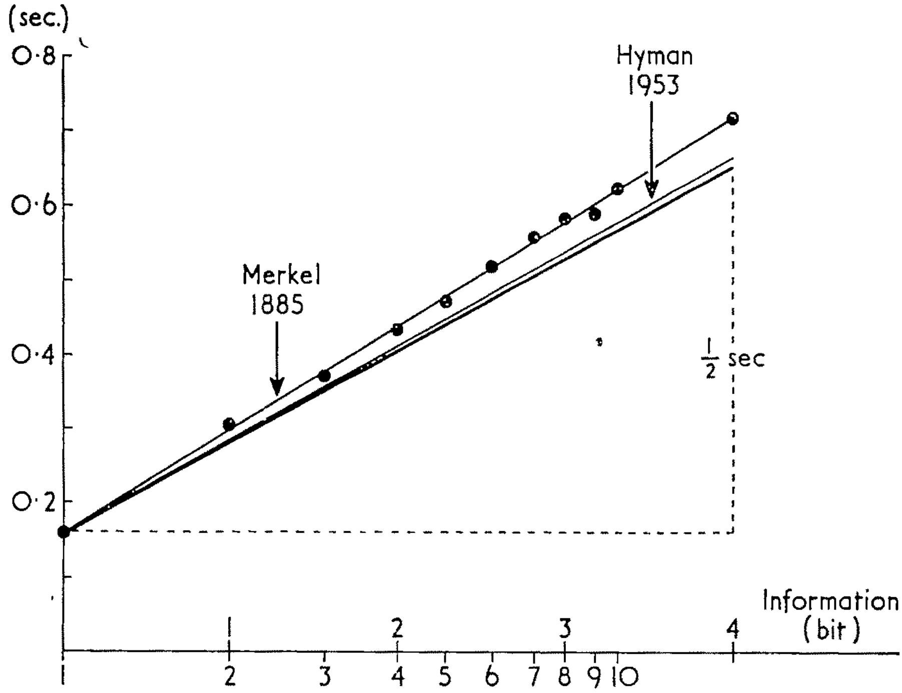


FIGURE 4.

Relation between reaction time in seconds and complexity of task, in bits. Data from Merker (1885) and Hyman (1953). (The Hyman data show results before and after practice) After Frank (1963).

V.—LEARNING AND INTELLIGENCE.

The theory that *learning* is basic to intellectual functioning is not necessarily antagonistic to a theory stressing speed ; within the more general speed theory we might expect that speedy learning would be characteristic of the bright, slow learning of the dull. In other words, learning would be one of the ' mental processes ' sub-divisions in Fig. 2. The early work of Woodrow (1946) was often considered to have disproved such an hypothesis, but his experiments were too simple altogether to throw much light on the problem ; it is not adequate to take subjects who are at different stages of mastery and practice on various types of tasks, who are differentially motivated towards thsee tasks, and who vary considerably with respect to the abilities involved in these tasks, and then to correlate speed of learning on these tasks with each other and with I.Q. Improved experimental and statistical methods have given more positive results regarding the relationship between I.Q. and learning (Stake, 1961 ; Duncanson, 1964).*

Another argument has often been presented, e.g., by Wechsler ; he has pointed out that a learning task such as ' memory span ' correlates poorly with the other tests in the W.A.I.S. and does not predict final total score well. Jensen (1964) has argued that this view is based on a neglect of the low reliability of the test as described by Wechsler ; this, in turn, can be raised to any height by simply lengthening the (very short) test, or by improving its design, or both. When correlations are corrected for attenuation, Jensen shows that digit span correlates ·75 with total I.Q., has a factor loading of ·8 on a general factor extracted from the Wechsler tests, is more culture-free than other tests, and can be shown to obey the Spearman-Brown prophecy formula, thus making it possible to increase its relability to any desired degree. The test can be made more predictive of I.Q. by measuring forward and backward span separately, rather than by throwing them together into one score ; apparently these two measures are not, in fact, highly correlated and should not be averaged but combined in some multiple correlation formula, if at all.

Jensen has used Digit Span and serial learning experiments of the traditional laboratory kind in an extensive investigation into personality determinants of invididual differences in these tests ; we shall return to this study later. Here it is relevant to mention that he found a multiple correlation of +0·76 between learning ability as so measured and college Grade Point Average, a measure of academic standing. When it is considered that this value was obtained in a relatively homogenous group of persons from the point of view of I.Q., and that this correlation is considerably higher than those usually reported with highly regarded I.Q. tests, then it may become apparent why I am suggesting here that we should take seriously the theory relating the concept of ' intelligence ' to learning efficiency and speed, and attempt, by means of laboratory studies such as those of Jensen and Roth, to investigate deductions from such an hypothesis. It seems reasonable to expect that such investigations are more likely to help in the elucidation of the nature of intellectual functioning than is the continued construction of I.Q. tests of a kind that has not materially altered in fifty years. And it is also possible that from the practical point of view, this method of procedure may result in tests and devices which enable us to give better predictions of school and university success than do existing tests.*

---

* An early study showing the close relation obtaining between intelligence, on the one hand, and learning/memory, on the other, was an investigation by Eysenck and Halstead (1945) of fifteen learning/memory tests ; these were found to be highly correlated with intelligence. A factorial analysis gave rise to a general factor of intelligence, leaving no residual evidence of any additional contribution by learning or memory.

As an example of the much increased possibility of psychological analysis opened by the use of laboratory methods in this field, consider Schonfield's (1965) study of memory changes with age. The general loss of ability of the aged to do I.Q. tests well has been known for a long time, as has their failure to acquire new skills and information, or to retain acquired material. These defects may be due either to a loss of ability to retrieve memories from storage, or to a deficiency in the storage system itself. By comparing recall and recognition scores on a learning task, Schonfield showed that recall was impaired in aged subjects, but recognition was not; he concluded that it was retrieval from memory storage which was at fault, rather than storage itself, thus suggesting that learning itself might be unimpaired with age. This experiment is cited, not because the results are definitive in any way, but because they illustrate well the approach suggested here; simple I.Q. testing cannot in the nature of things do any more than reveal the existence of a deficit, but in order to reveal the precise psychological nature of the intellectual deficit in question more experimental methods are required.

VI.—LEARNING AND PERSONALITY.

In our discussion of Furneaux's contribution, we found that of his three components of intellectual functioning, only one (speed) was cognitive, while two (persistence and the error-checking mechanism) seemed more orectic in origin, and likely to be related to personality. Most workers in the field of intelligence testing disregard personality factors altogether, but this is almost certainly a mistake. There are several experiments which bring out fairly clearly the importance of personality factors such as neuroticism and extraversion/introversion in the measurement of intelligence, and much of our work has centred on this aspect. Consider first of all the simple learning experiments which we have just discussed; here one can perhaps expect personality to play little if any role. This, however, is not so, and it may be interesting to speculate about the kind of relation which one might expect to find. We may with advantage begin by considering the well-known experiments of Kleinsmith and Kaplan (1963). These authors argued, briefly, that learning is mediated by a *consolidation process* which takes place after the learned material has been registered, but before it is transferred into permanent memory storage. Consolidation is a function of the state of arousal of the organism; the greater the arousal, the longer and more efficient the consolidation, so that higher arousal leads to better memory in the long run. However, while consolidation is proceeding, it interferes with recall, so that while the consolidation process is going on the highly aroused organism is at a disadvantage. Kleinsmith and Kaplan tested their theory by measuring the amount of arousal (G.S.R. reaction) produced by different paired stimuli; for each subject they then picked the most arousing and the least arousing stimulus pairs and had the subject remember the paired stimulus after presentation of the original stimulus.

---

* One interesting possibility which is suggested by Jensen's work relates to his finding that serial learning tasks and paired associate learning tasks both correlate with I.Q., but not with each other. In view of the dependence of paired associate learning on verbal mediation, in contrast to the rote character of serial learning, it seems possible to regard serial learning as the prototype of Cattell's 'fluid' ability, and paired associates learning as the prototype of his 'crystallized' ability (1964). If this suggestion has any value, it may show the way to the construction of a battery of tests less dependent on cultural factors and training than are most existing I.Q. tests.

Recall was arranged at different times after original learning for different groups of subjects, and Fig. 5 shows the results ; it will be seen that as expected high arousal words are poorly remembered immediately after learning, but show very marked reminiscence effects, while low arousal words are well remembered immediately after learning, but fade out quickly. There is little doubt of the reality of this phenomenon, which has since been demonstrated several times.

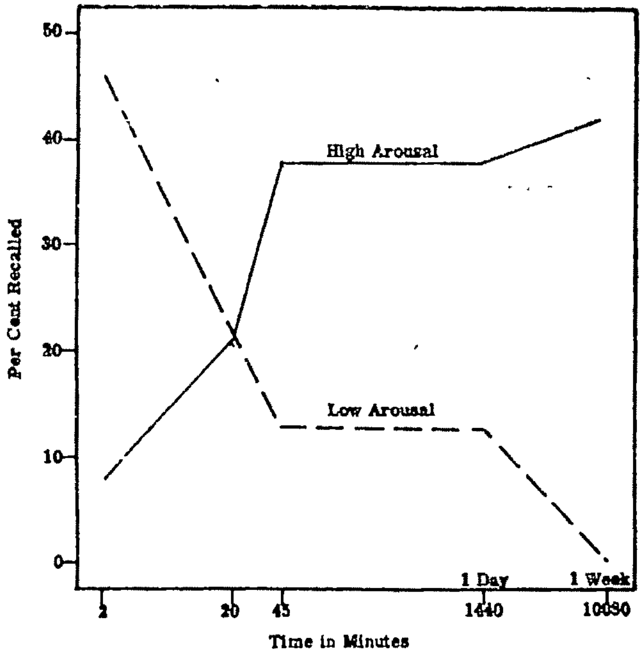


FIGURE 5.
Differential recall of paired associates as a function of arousal level (Kleinsmith and Kaplan, 1963).

In this experiment stimuli were measured and grouped according to their arousing qualities. It is equally possible to group subjects according to their arousability, and I have argued that introverts are characterised by high arousal, extraverts by poor arousal (Eysenck, 1963, 1967). If this theory is along the right lines, we would expect extraverts to behave in the manner of the low arousal words in Fig. 5, and introverts in the manner of the high arousal words. In other words, for short recall times, extraverts should be superior, while for long recall times introverts should be superior. There are about half-a-dozen experiments in the literature demonstrating the superiority of extraverts over short-term intervals, including the work of Jensen already mentioned ; these have been summarized elsewhere (Eysenck, 1967), and all that need be said here is that results are in good agreement with prediction. Some unpublished work on pursuit rotor reminiscence also supports the prediction of better learning for introverts after long rest intervals.

A specially designed experiment by McLaughlin and Eysenck was undertaken to test, in addition to the hypothesis stated above, a further one relating to the personality dimension of neuroticism, which we may regard as associated with drive (Spence, 1964). Subjects were tested on either an easy list of seven pairs of nonsense syllables, or on a difficult list, difficulty being manipulated through degree of response similarity. It was predicted that in accordance with the Yerkes-Dodson law the optimum drive level for the easy list would be higher than that for the difficult list, and it was further assumed that N subjects (high scorers on the N scale of the E.P.I.) would be characterized by higher drive than S subjects (stable, low scorers on the N scale of the E.P.I.). Extraverts, as already explained, were regarded as low in arousal, introverts as high. There are thus four groups of subjects, which, in order of drive, would be (from low to high) : stable extraverts ; neurotic extraverts and stable introverts ; neurotic introverts. (No prediction could be made about the position of the two intermediate groups relative to each other.) The results of the experiment are shown in Fig. 6 ; extraverts, as predicted, are significantly superior to introverts, and the optimum performance level of drive is shifted towards the low end as we go from the easy to the difficult list, thus shifting the SE group up and the NE group down. (The figures in the diagram refer to number of errors to criterion.)

If introverts, as hypothesized, are characterized by a more efficient consolidation process, due to their greater cortical arousal, then we should be able to predict that they should be superior to extraverts with respect to acquired knowledge. As an example, we may take vocabulary scores, which are clearly the product of learning, and which usually correlate very highly with other I.Q. tests. Eysenck (1947) has reported personality differences between 250 neurotic male soldiers whose Matrices scores were much superior to their Mill Hill Vocabulary scores, and 290 male soldiers whose scores showed a similar difference in the opposite direction ; he also studied 200 and 140 neurotic women soldiers showing similar differences. In both sex groups those subjects whose vocabulary was relatively good showed dysthmic (introverted) symptoms, while those whose vocabulary was relatively poor showed hysteric (extraverted) symptoms. Farley (unpublished) has carried out a study of forty-seven normal subjects in which he found a substantial positive correlation ($r=+0{\cdot}48$) between introversion and vocabulary. This is of course in line with the alleged ' bookish ' character of the typical introvert. There was no such correlation between Introversion and Raven's Matrices.

It is possible to go further than this and argue that introverts should do rather better at school and university because of this superiority in consolidation of learned material ; there is much evidence to indicate that such a prediction may be along the right lines (Furneaux, 1962 ; Lynn and Gordon, 1961; Savage, 1962 ; Bendig, 1958, 1960 ; Otto, 1965 ; Otto and Fredricks, 1963 ; Child, 1964 ; Ranking, 1963a, 1963b). Not all the results are favourable, but the overall impression is certainly in accordance with expectation. It might be suggested that some form of zone analysis (Eysenck, 1966) which included the N variable as well as the E variable would throw much needed light on these relationships. It should be added that the results do not so much support the hypothesis, as rather fail to disprove it. There are so many alternative hypotheses to account for the finding that not too much should be read into the data.

VII.—INTELLIGENCE AND PERSONALITY.

It will be clear from this discussion that personality features such as neuroticism and extraversion-introversion interact with learning in complex though meaningful ways, and that great care has to be taken in the design of

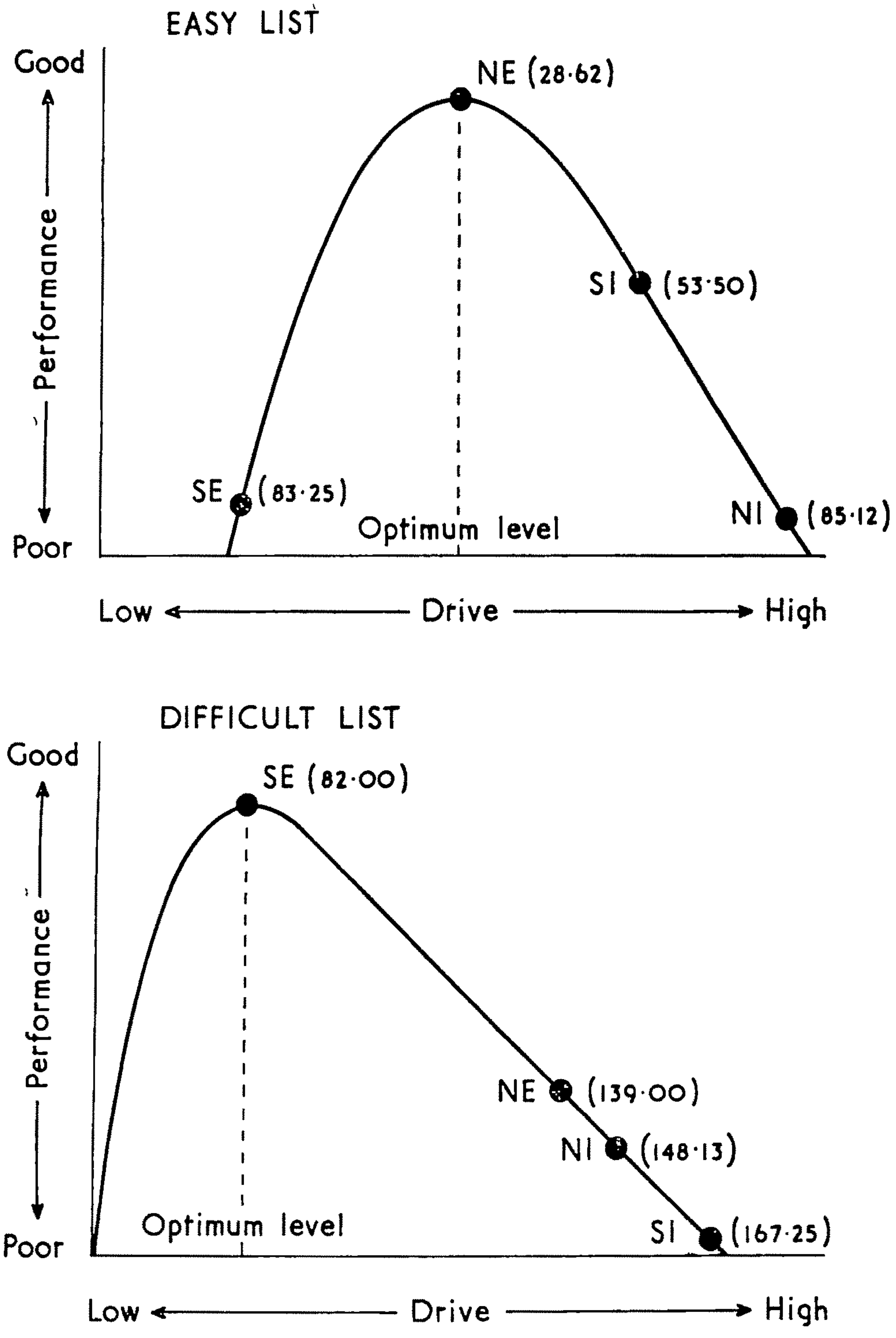


Figure 6.

Performance of stable extraverts, neurotic extraverts, stable introverts and neurotic introverts on easy and difficult paired associate learning tasks (R. J. McLaughlin and H. J. Eysenck, unpublished data).

experiment not to fall foul of the complex laws relating performance to personality.* It might be objected that such relations only obtain when laboratory learning tasks are used, but that they fail to appear when orthodox intelligence tests are employed. This is not so. One of the earliest findings relating to extraversion/introversion was that extraverts opt for speed, introverts for accuracy, when there is the possibility of a choice in the carrying out of an experimental task (Eysenck, 1947), and we would expect this difference to appear in relation to intelligence tests also. Jensen (1964) correlated extraversion scores on the E.P.I. with time spent on the Progressive Matrices test and found a significant correlation of —0·46; in other words, extraverts carried out the task more quickly. They also made more errors, but this trend was not significant. Farley (1966) applied the Nufferno test individually to thirty Ss, divided on the basis of their E.P.I. scores into ten extraverts, ten ambiverts and ten introverts. The mean log speed scores on all problems correctly solved for the groups were respectively: ·78, ·88 and ·93. This monotonic increase in solution time with introversion was fully significant by analysis of variance. Other examples of this relation between speed and extraversion are given elsewhere (Eysenck, 1967); there seems little doubt about its reality.

Farley (1966) also discovered a significant relation with neuroticism, but as might have been expected (Payne, 1960) this showed a non-linear trend, subjects with average scores being superior to those with high or low N scores. Lynn and Gordon (1961) have also published a study showing a similar trend; they used the Progressive Matrices test. The rationale underlying the prediction of a curvilinear relationship in this context derives, of course, from the Yerkes-Dodson law; it is believed that the optimum drive level for complex and difficult tasks like those involved in an intelligence test lies below the high level reached by high N subjects, and above that reached by low N subjects. The general drive level of the group tested is, of course, quite critical in this connection, and it must be emphasized that unless this can be specified or measured, predictions will not always be fulfilled. Changes in difficulty level of the items, changes in the importance the result of the test assumes in the eyes of the subjects, and changes in the motivational value of the instructions may all lead to a general shift in the drive level of the subjects which may displace the optimum level in either direction. It would seem useful in tests of this prediction to have separate measures of drive, or of arousal, against which performance could be plotted (Eysenck, 1967): without such direct measures the subjects' N score may often be difficult to interpret, giving us essentially merely a measure of their *probabability* of responding with antonomic activation to an anxiety-producing situation. If the situation is not perceived as anxiety-producing by the subjects, then differences in N cease to matter. This line of argument has led to a better understanding of the conditions under which N correlates with eyeblink conditioning (Eysenck, 1967), and it may be used to design experiments explicitly aimed at increasing the correlation posited.

This dependence of results on precise control of parameter values can also be illustrated by some recent unpublished experiments undertaken by M. Berger. We have noted that extraverts are faster and make more errors when conditions are such that the test is administered without stress on speed; in other words, when no explicit instructions are given emphasizing speed, extraverts opt for speed and neglect accuracy, whillle introverts opt for accuracy and go slow. These are response styles well familiar from other types of activity (Eysenck, 1960). What would we expect to happen when stress was placed, explicitly and implicitly, on speed of problem solving activity? Let us return to Fig. 6, in which we postulated that stable extraverts would have low drive level. neurotic introverts high drive level, with the other two groups (stable introverts and neurotic extraverts) intermediate. Given the specific stress on *speed* as the proper index of performance, we would expect the low-drive stable extraverts to have the slowest speed, and the neurotic introverts the highest, with the other two groups intermediate; we might also expect that the neurotic introverts would produce more errors in order to make up for the excessive speed shown.

---

* The common belief that incentives and higher or depressed motivation generally do not affect intelligence test performance (Eysenck, 1944) may be mistaken; it is conceivable that here too we find the curvilinear Yerkes-Dodson relation, so that *overall* failure to find significant motivation and incentives may be due to compensating positive and negative effects of increased motivation on different types of subject Some form of interaction terms should be included in analyses of this type, and by preference this should take the form of zone analysis. (Eysenck, 1966.)

Berger tested twenty-one 13-year-old school children in each of the four personality groups; the groups were equated for age, sex and intelligence, using their 11+ records for this purpose. Fifty problems were presented for solution individually, followed by a rest, and finally by another set of thirty problems. Each problem was shown to the child on a screen, with numbered alternative solution; having selected the correct solution, the child pressed a numbered button, which activated a time switch, thus recording solution latency, and also caused the projector to project the next problem on to the screen. Instructions emphasized speed of working, and the whole experimental set-up added to this impression; furthermore, the disappearance of the problem after the button had been pressed eliminated the possibility of checking the correctness of the answer. Figure 7 shows the results of the first fifty items; the next thirty showed similar results. The Figure is arranged in the form of a cumulative time record, with time arranged along the ordinate and the problems, 1 to 50, along the abscissa. It will be clear that the stable extraverts are much the slowest, the neurotic introverts much the fastest, with the other two groups intermediate; these differences are highly significant. It was also found, at a high level of significance, that neurotic introverts compensated for their speed by making more errors than the other groups. Thus, the Yerkes-Dodson law appears to be working here very much as it did in the case of the McLaughlin-Eysenck experiment: the low drive SE group does poorly because it is so slow, the high drive NI group does poorly because it makes too many errors, and the intermediate NE and SI groups do best because they work at an optimum level of motivation.

VIII.—FLUENCY AND INHIBITION.

This study illustrates the value of applying theories and laws from general and experimental psychology to intelligence testing. Another example may serve the same function. From the point of view of the experimental psychologist, a typical intelligence test is a good example of a task undertaken in the condition of massed practice; we would, therefore, expect it to generate reactive inhibition. Extraverts generate such inhibition more strongly and more quickly than do introverts (Eysenck, 1957, 1967), and consequently we would expect that when groups of extraverts and introverts are matched for performance during the earlier part of an intelligence test, then they will diverge towards the latter part, with the introverts superior in performance. Another way of saying the same thing would be to regard an intelligence test as a vigilance test, and use the well-known fact that introverts preserve vigilance better than extraverts

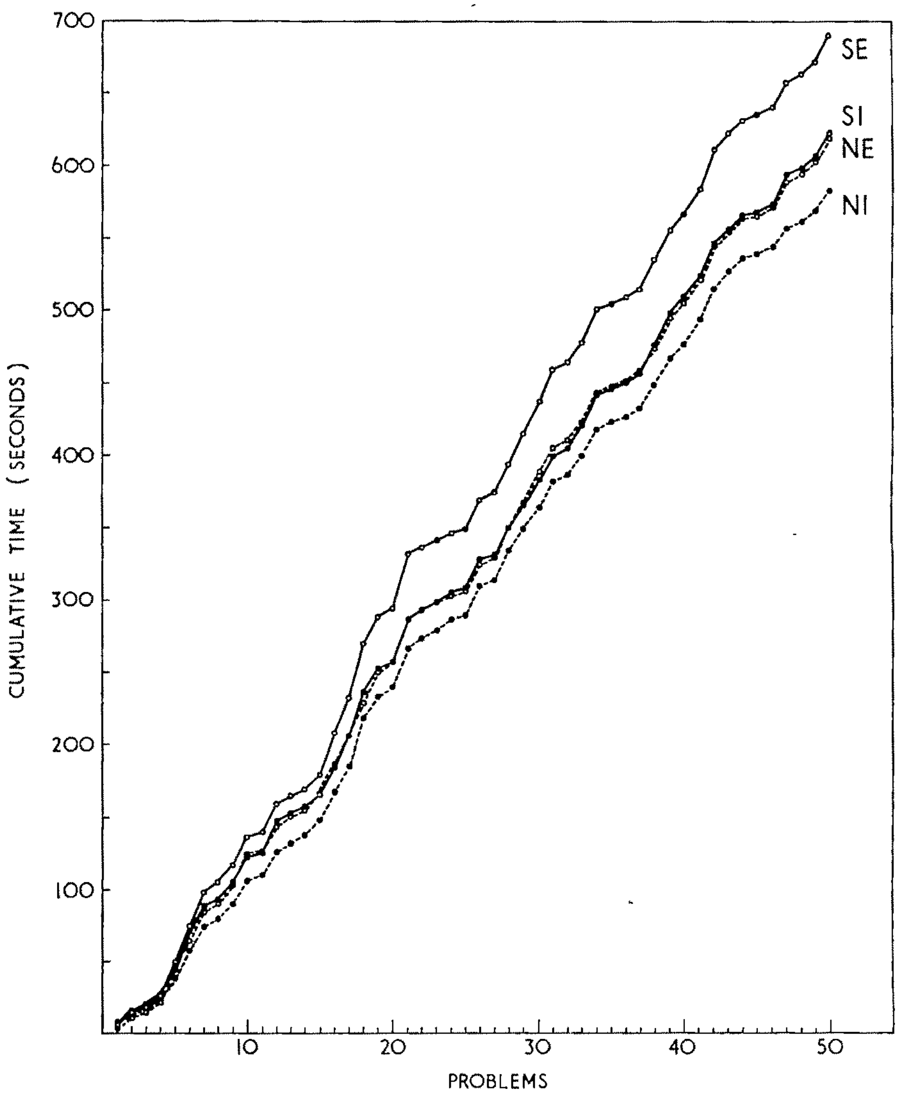


FIGURE 7.
Solution times of stable extraverts, neurotic extraverts, stable introverts and neurotic introverts on fifty intelligence problems, timed separately and cumulated. (M. Berger, unpublished data.)

to predict their better performance towards the end. Eysenck (1959) has reported such an experiment, in which he used sixty items from the Morrisby Compound Series test, individually but unobtrusively timed. Using speed of correct solutions, it was found that on the first forty-five problems introverts were slower than extraverts, but on the last fifteen items, the two groups reversed position and the extraverts were now the slower. On the last fifteen items, it was also found that the extraverts gave up more easily. It would thus seem true to say that extraverts do show the predicted decline in performance during the latter part of their performance on a typical test of intelligence, administered as far as the subjects were concerned in the usual manner, and without any special stress on speed. This experiment, taken in connection with the others already quoted, leaves little doubt that personality plays an important part in intelligence test performance, and that its influence has hitherto been very much under-estimated.

Personality factors interact with intelligence test performance in many ways, and neglect of these factors may easily lead to quite incorrect conclusions. As an example, we may, perhaps, take the large body of work recently done on convergent and divergent types of tests (Hudson, 1966). In studies of this kind, candidates good on divergent tests are often called ' creative,' and the argument is sometimes extended to other desirable qualities of intellect, such as ' originality ' (Barron, 1963 ; Taylor and Barron, 1963). In fact, divergent tests are by no means new ; under the title of ' fluency ' tests they were among the early discoveries of the London school, and a typical set of such tests is reprinted in Cattell's (1936) *Manual of Mental Tests.* Tests of this kind were found to be correlated with extraversion (Eysenck, 1960) and Spearman (1926) already pointed out that this particular factor " has proved to be the main ground on which persons become reputed for ' quickness ' or for ' originality '." Hudson's work supports some such interpretation quite strongly ; ' divergent ' schoolboys, as compared with ' convergent ' ones, are more fluent, make more errors on orthodox tests, are emotionally more forthcoming, are more sociable, and prefer ' arts ' to ' science ' subjects—all characteristics of extraverts as compared with introverts. There is, in fact (as Hudson acknowledges) no evidence to show that ' divergent ' boys are more creative than ' convergent ' ones ; as he points out, one can be ' creative ' in different ways. All that we seem to be dealing with in this distinction would seem to be a kind of response set or ' style'; it is, perhaps, unusual to apply this concept in relation to intelligence tests, but it applies here probably more than in relation to personality inventories.

IX.—THE LIMITATION OF PSYCHOMETRY.

These various ways in which personality and intelligence testing interact do not by any means exhaust the available evidence. Factor analysts usually assume, without proof, that groups which do not differ in performance on a group of tests will also not differ in factorial solution. Lienert (1963) showed that this assumption is, in fact, erroneous ; children high and low on N, respectively, do not produce identical correlation matrices or factors, when administered sets of intelligence tests, nor do the two groups even agree in the number of factors produced. As Eysenck and White (1964) have shown in a re-analysis of the data, " the stable group has a more clearly marked structure in the cognitive test field than has the labile group." (It has also been found that students differing in intelligence do not have identical factor patterns on personality questionnaires ; the evidence is presented by Shure and Rogers, 1963). It is not unlikely that some of the observed differences in factor structure are connected with the intellectual response styles which we have found to be characteristic of different personality groups, but at present there is no evidence to indicate precisely how this may have come about. Much further work is clearly required before we can be sure of our facts in this complex field.

All that has been said in this paper is only suggestive, and I do not in any way believe that the hypotheses stated, and sometimes supported by experimental data, are at the moment anything but guideposts pointing in the direction of interesting and important factors which will almost certainly have a bearing on the proper measurement of intelligence. We have noted four stages in the development of intelligence tests ; it is the main purpose of this paper to suggest the importance of starting out on a fifth stage of intelligence assessment, a new stage based on theoretical and experimental work, and not divorced from the main body of academic psychology. Psychometrics and factor analysis have important contributions to make, but they can do so only in conjunction with other disciplines, not by 'going it alone.' What is required is clearly an integration of intelligence testing with the main stream of academic psychology, and a more determined experimental and laboratory approach to the problems raised by the various theories of intellectual functioning. Some obvious suggestions emerge from the inevitably somewhat rambling and unco-ordinated discussion of this paper. (1) Analysis of performance should always take into account individual items, rather than tests, i.e., averages taken over what may be, and usually are, non-homogeneous sets of items. Such analysis should be made in terms of latencies, i.e., speed of individual item solution, as well as of errors, persistence before abandoning items, and other similar differential indicators of response style. (2) Investigators should pay more attention to laboratory studies of learning and memory functions, of speed of information processing, and other experimental measures in the testing of specific hypotheses regarding the nature of intellectual functioning. Analysis of intelligence tests of the orthodox kind raises problems, but cannot in the nature of things go very far towards answering them. (3) Investigators should experiment with variations in experimental parameters, such as rest pauses, time from end of learning to recall, rate of presentation, degree of motivation, etc., in an effort to support or disprove specific theoretical predictions regarding the process of learning and problem solving. (4) Personality variables, such as stability-neuroticism and extraversion-introversion, should always be included in experimental studies of intellectual functioning, because of their proven value in mediating predictions and their interaction potential in all types of learning and performance tasks. Vigorous research along these lines carries with it the promise that notions such as intelligence, I.Q., ability and factor will cease to be regarded as poor relations, and will return to the eminent and successful status they held before the war; it also furnishes the only means of making these concepts scientifically meaningful, academically respectable, and practically more useful.

X.—References.

Barron, F. (1963). *Creativity and Psychological Health.* Princeton : Nostrand.

Bendig, A. W. (1958). Extraversion, neuroticism and verbal ability measures. *J. Consult. Psychol.*, 22, 464.

Bendig, A. W. (1960). Extraversion, neuroticism and student achievement in introductory psychology. *J. Educ. Res.* 53 263-267.

Binet, A , and Simon, R. (1905). Méthodes nouvelles pour le diagnostic du niveau intellectuel des anormaux. *Annee Psychol.*, 11, 191-244.

Cattell, J. McK. (1890). Mental tests and measurements. *Mind*, 15, 373-380.

Cattell, R. B. (1936). *A Guide to Mental Testing.* London : University of London Press.

Cattell, R. B. (1964). Fluid and crystallized abilities. In R. B. Cattell, *Personality and Social Psychology.* San Diego : R. R. Knapp.

Child, D. (1964). The relationships between introversion extraversion, neuroticism and performance in school examinations. *Brit. J. Educ. Psychol.*, 34, 187-196.

Duncanson, J. P. (1964). *Intelligence and the Ability to Learn.* Princeton, N.J. : Educational Testing Service.

Eysenck, H. J. (1939). Primary mental abilities. *Brit. J. Educ. Psychol.*, 9, 270-275.

Eysenck, H. J. (1944). The effect of incentives on neurotics, and the variability of neurotics as compared with normals. *Brit. J. Med. Psychol.*, 20, 100-103.

Eysenck, H. J. (1947). *Dimensions of Personality.* London : Routledge and Kegan Paul.

Eysenck, H. J. (1953). *Uses and Abuses of Psychology.* London : Pelican.

Eysenck, H. J. (1957). *The Dynamics of Anxiety and Hysteria.* London : Routledge and Kegan Paul.

Eysenck, H. J. (1959). Personality and problem solving. *Psychol. Rep.*, 5, 592.

Eysenck, H. J. (1960). *The Structure of Human Personality.* London : Methuen.

Eysenck, H. J. (1963). The biological basis of personality. *Nature*, 199, 1031-1034.

Eysenck, H. J. (1966). Personality and experimental psychology. *Bull. Brit. Psychol. Soc.*, 62, 1-28.

Eysenck, H. J. (1967). *The Biological Basis of Personality.* New York : C. C. Thomas.

Eysenck, H. J., and Halstead, H. (1945). The memory function. *Amer. J. Psychiatry*, 102, 174-180.

Eysenck, H. J., and White, P. O. (1964). Personality and the measurement of intelligence. *Brit. J. Educ. Psychol.*, 34, 197-202.

Farley, F. H. (1966). Individual differences in solution time in error-free problem solving *Brit. J. Soc. Clin. Psychol.*, 5. To appear.

Frank, H. (1963). Informations psychologie and nachrichtentechnik. In Weiner, N., and Schade, J. P. *Progress in Brain Research, Vol. 2, Nerve, Brain and Memory Models.* Elsevier Publishing Co., Amsterdam, 79-96.

Furneaux, W. D. (1960). Intellectual abilities and problem solving behaviour. In H. J. Eysenck (Ed.) *Handbook of Abnormal Psychology.* London : Pitman.

Furneaux, W. D. (1962). The psychologist and the university. *Univ. Quart.*, 17, 33-47.

Galton, F. (1883). *Inquiries Into Human Faculty and its Development.* London : Macmillan.

Guilford, J. P. (1966). Intelligence : 1965 model. *Amer. Psychol.*, 21, 20-26.

Hick, W. (1952). On the rate of gain of information. *Quart. J. Exp. Psychol.*, 4, 11-26.

Hudson, L. (1966). *Contrary Imaginations.* London : Methuen.

Hyman, R. (1953). Stimulus information as a determinant of recation time. *J. Exp. Psychol.*, 45, 188-196.

Jensen, A. R. (1964). *Individual Differences in Learning : Interference Factors.* U.S. Depot. of Health, Educatisn and Welfare. Co-op. Project No. 1867.

Kleinsmith, L. J., and Kaplan, S. (1963). Paired-associate learning as a function and interpolated interval. *J. Exp. Psychol.*, 65, 190-193.

Lienert, A. A. (1963). Die Faktorenstrukter der intelligenz als Funktion des Neurotizismus. *Zhstr. F. Exp. Angew. Psychol.*, 10, 140-159.

Lynn, R., and Gordon, I. E. (1961). The relation of neuroticism and extraversion to intelligence and educational attainment. *Brit. J. Educ. Psychol.*, 31, 194-203.

McNemar, R. (1964). Lost . our intelligence ?: Why ? *Amer. Psychol.*, 19, 871-882.

Merkel, J. (1885). Die zeitlichen Verhältnisse der Willenstatigkeit. *Philos, Studien*, 2, 73-127.

Otto, W. (1965). Inhibition potential in good and poor achievers. *J. Educ. Psychol.*, 56, 200-207.

Otto, W., and Fredricks, R. C. (1963). Relationship of reactive inhibition to reading skill achievement. *Educ. Psychol.*, 54, 227, 230.

Payne, R. W. (1960). Cognitive abnormalities. In H. J. Eysenck (Ed.) *Handbook of Abnormal Psychology.* London : Pitman.

Rankin, G. F. (1963a). *Reading Test Performance of Introverts and Extraverts.* Milwaukee 12th Yearbook of Nat. Reading. Conf.

Rankin, G. F. (1963). Reading test reliability and validity as function of introversion. extraversion. *J. Devel. Reading*, 6, 106-117.

Eysenck, H. J. (1967). *The Biological Basis of Personality.*

Roth, E. Die Geschwindigkeit der Verarbeitung von Information und ihr Zusammenharg mit intelligenz. *Atschr F. Exp. Angew. Psychol.*, 11, 616-622.

Savage, R. D. (1962). Personality factors and academic performance. *Brit. J. Educ. Pschol.*, 32, 251-252.

Schmidtke, H. (1961). Zur Frage der informationstheoretischen Analyse von Wahlreaktionsexperimenten. *Psychol., Forschung*, 26, 157-178.

Schonfield, D. (1965). Memory changes with age. *Nature*, 208, 918.

Shure, G. H., and Rogers, M. S. (1963). Personality factor stability for three ability levels. *J. Psychol.*, 55, 445-456.

Spearman, C. (1904). " General intelligence " objectivity determined and measured. *Amer. J. Psychol.*, 15, 201-293.

Spearman, C. (1923). *The Nature of " Intelligence " and the Principles of Cognition.* London : Macmillan.

Spearman, C. (1927). *The Abilities of Man.* London : Methuen.

SPENCE, K. W. (1964). Anxiety (drive) level and performance in eyelid conditioning. PSYCHOL., BULL., 61, 129-139.
STAKE, R. E. (1961). Learning parameters, aptitudes and achievements. *Psychometric Mon.*, No. 9.
TAYLOR, C. W., and BARROW, F. (Eds.) (1963). *Scientific Creativity*. London : Wiley.
THORNDIKE, E. L. (1926). *The Measurement of Intelligence*. New York : Teachers' College Columbia University.
THURSTONE, L. L. (1926). *The Nature of Intelligence*. New York : Harcourt, Brace and World.
VERNON, P. E. (1965). Ability factors and environmental influences. *Amer. Psychol.*, 20, 273-733.
WOODROW, H. (1946). The ability to learn. *Psychol. Rev.*, 53, 147-158.

(*Manuscript received* 21*st July*, 1966)

CONCEPTS OF TIME IN SECONDARY SCHOOL CHILDREN OF ABOVE AVERAGE I.Q.

By K. W. ROGERS

(*University of Sheffield, Institute of Education*)*

SUMMARY. Previous research tends to equate the concept of time with the concept of historical time so that the concept has not been fully explored. It was proposed that an integral concept of time exists and a test was devised to measure it. This test, together with a test of the historical concept, was presented to four consecutive age groups in the academic streams of a mixed comprehensive school. Analysed for variance, the scores on the test of the integral concept showed an improvement, significant at the ·01 level, between the second and third groups, whose average ages were 13·6 and 14·6, respectively. An improvement on the test of the historical concept, significant at the ·05 level, took place between the first and second groups, whose average ages were 12·6 and 13·6, respectively. The inference drawn was that mastery of the historical concept does not imply mastery of the integral concept and that a clear distinction may be made between the two.

Boys proved significantly (·01 level) superior to girls on the test of the integral concept, a fact which may relate to the apparent superiority of boys in the mastery of abstract ideas in general.

I.—INTRODUCTION.

THE first paragraph of any recent article dealing with temporal concepts will frequently be found to contain a reference to the lack of published work on the subject. Smythe and Goldstone report upon " . . . the dearth of experimental data concerning the development of temporal concepts of children" (1957, p. 49), while Wallace and Rabin refer to " . . . the relatively uncrystallised state of the subject matter " (1960, p. 213). As late as March, 1964, Ed. E. Johnson writes that " . . . little effort has been made to understand concepts of time as compared with the extensive study of general intelligence, special abilities and the senses " (1964, p. 377). Indeed, when the general field of the literature since Sturt (1922) is surveyed, there is little evidence of progress in either the test methods used or the conclusions arrived at. Johnson himself used a test originally designed by Buck, in 1946, which itself was not greatly different from those used by Sturt. In fact, many works on time concepts of recent origin seem to take the form of reflections upon past research in an effort to find fresh growing points (cf. Wallace and Rabin 1960, Jahoda 1963). Considerably more work has been done in the related fields : the perception or judgment of periods of short duration and the subjective distortion of time by neurotic and mentally disordered subjects. The reason for this, perhaps, is that testing in these fields would seem to present fewer problems. Some idea of the quantity of work on the psychology of time in general, and the proportion of that work which has concerned itself with the concept of time, may be obtained from the excellent bibliography appended to the article by Wallace and Rabin (1960).

Those writers who have attempted to deal with the nature of the time concept as such, rather than the earlier stages of the development of time notions, seem very largely to have confined themselves to the concept of historical time, or at least to have included so much historical material in their tests as to achieve this effect (e.g., Pistor 1939, 1940 ; Friedman 1943, 1944, 1945 ;

---

* Now at North Buckingham College of Education.

Rawlins 1956). The assumption made by such authors would seem to be that the concept of historical time represents the ultimate in the elaboration of the concept of time normally achieved. It is possible, however, that the historical concept is merely one aspect of a wider concept in which all the various facets of time, as they are presented to us, are integrated.

The essential character of a concept is its generality. Lovell defines a concept as " a generalisation about data which are related ; it enables one to respond to, or think about, specific stimuli or percepts in a particular way. Hence, a concept is exercised as an act of judgment " (Lovell, 1960, p. 13). The fuller the generalisation of the concept, the greater the flexibility of the individual's responses to stimuli which involve it, and the better his judgment in problems relating to the concept. Now stimuli involving time are continuously present in our everyday life : we reflect upon recent or less recent events, we forecast durations, long and short, we read books or newspapers which require us to imagine events in the historical past, or for that matter, in the distant future, if one favours science fiction. Many of these forms of behaviour in respect of time involve the use of measuring instruments such as the clock or the calendar and have been acquired at a relatively early age : almost certainly by 10 years at least. Investigations have shown that such forms of behaviour, and the facility in the use of measuring instruments seem to appear in a relatively uniform sequence and at the same relative time in the life of any child (cf. Ames, 1946 ; Schachter, Symonds and Bernstein, 1955). It would be dangerous, however, to suppose that the ability to use measuring instruments in respect of time implies that the child has fully acquired the concept of time, for, as Jahoda pointed out : " Knowledge of the verbal symbol is no guarantee that the corresponding concept has been fully assimilated " (1963, p. 88).

Whatever the form which the early development of temporal notions may take, it is reasonably safe to assume that, at some stage, the individual attains the ability to generalise about such notions. That is to say, he perceives the common element in each of the various notions or modes of behaviour and so arrives at the concept of time itself as an abstraction independent of the spatial framework within which time has hitherto been appreciated. In Piaget's terms, a stage is reached at which concrete operations in respect of time merge into formal operations.

That we in our culture tend to think of time as regular, measured and immutable in the speed of its passing, is demonstrated by the difficulty which most of us have in trying to comprehend Einstein's theories about the relationship between space and time and to accept the proposition that it is possible to travel in space so as to leave the time continuum of Earth. The elaborations of the concept of time undertaken by mathematicians, however, need not concern us here, for the purpose of this discussion is merely to delimit the concept of time as it exists in the mind of the normal individual going about his business : the norm which the culture of our society demands of its members. We expect the normal adult to be able to read, to be capable of elementary arithmetic, to possess certain social attitudes, and similarly, to be able to tell the time and imagine the past and future. To extend the concept further may be the metier of certain specialists but it is not required of us by society. As Paul Fraisse says of the scientists' attempt to define universal time : " . . . il n'affecte pas la vie psychologique quotidienne, . . . " (1957, p. 9). For most adults there would seem to exist a time which, to them, is real and universal. We might appropriately refer to this as Newtonian time. Perhaps the most common way of defining this time is as a fourth dimension, as essential to the existence of matter as the other three. Where it differs from them is that it has no existence for us independently of them. As Piaget says, space and time are indissociable: "... l'espace est un instant pris sur le temps et le temps est l'espace en mouvement" (1946, p. 2). Where we can make abstractions of space in geometrical constructions, without reference to the dimension of time, we are unable to do the opposite. Thus, Newtonian time may exist as a concept in the adult mind, but it is only accessible to the senses in terms of space: It is this inaccessibility of time which makes the concept a late developer compared with other concepts, a fact upon which all writers on the subject concur.

The concept of historical time, since it deals with time beyond the individual's personal temporal horizon, must be described as a formal operation, but it is, nevertheless, patently an abstraction derived from only one aspect of time. It is conceivable that one might be able to think in terms of historical time without possessing a full and integrated concept of Newtonian time. If this be so, then a distinction can be drawn between the two forms of time concept and we would be justified in describing the latter as the whole or integral concept of time. Furthermore, the inference could be drawn that the integral concept must include the historical concept and that, in their order of appearance in the process of development, the latter must precede the former.

Since it is not directly accessible to the senses, a good deal of subjective distortion of time is usual in individual behaviour (cf. Bergler and Roheim, 1946; Clausen, 1950). This fact has tended to cloud the issue in much of the research into time concepts. In educational research. we must be concerned with the commonly held concept and not with the behaviour of individuals in relation to time. In schools we would like to know what concepts are and how they develop so that we may perhaps learn how to encourage their growth. Whether a certain individual finds time passing slowly or quickly under certain circumstances is not often relevant to the teaching of history, French or physics, but whether he understands precisely what the time dimension is and how we may help him to comprehend it, can be of vital importance in the teaching of many subjects. This is not to suggest that subjective time is not a fit subject for psychological study or that teachers are not concerned with the personalities of their pupils.

This enquiry attempts to provide some evidence for the separate existence of the integral concept of time, as it has been defined above, and to draw the distinction between the integral and the historical concepts in terms of developmental sequence.

II.—METHOD.

Two tests of the historical concept and of the integral concept of time, were presented to four consecutive age groups in a mixed, comprehensive secondary school. Analysis of variance procedures were then applied to show the significance of difference between the means of the four age groups, pairing each group with the adjacent group and with each other group. The significance of the difference between the mean scores of boys and girls throughout the four groups was obtained by the same method. The correlation between the two tests was also calculated.

The test of the integral concept contained eighteen questions of various types, all awarded a single mark, or, if the question were subdivided, a mark for each subdivision. The maximum was 24. In order not to favour logico-mathematical or verbal ability, equal weighting within the test was given to questions which appeared to be of either type. Some questions, for example, were simple arithmetical calculations involving time while others were tests of the vocabulary of time words. Still other questions were concerned with historical chronology. In short, since time is a dimension of all things and is accessible through many different kinds of experience, it seemed that the best way to delimit the integral concept was to examine it from as many different angles as possible. Thus, the main criterion in framing the test needed to be the exploration of the ability to think in terms of time in a wide variety of contexts, which necessitated making the test items as variable as possible in their form while keeping, as their common content, the time dimension. The greater the subject's score, the greater, it was presumed, would be his grasp of the generalised concept of time.

Taken as a whole, the test can have little more than face validity since no previous investigation has employed this combination of test items, but the individual items have, in many cases, been used before, if not in these exact forms, and all of them have a manifest time content.

Certain authors place the beginning of the time concept in its adult form at about 11 years (cf. Sturt, 1922 ; Hunter, 1934 ; Bradley, 1947). Thereafter, a steady improvement takes place, according to such writers, until the full adult concept is reached. Friedman (1945) implies that this improvement continues to 17 or 18 years while Hunter (1934) suggests that it continues until at least 16. As we have already observed, such findings have been produced by tests which have tended to concentrate upon aspects of the historical concept, coupled, in some cases, with tests of the ability to comprehend the normally used time words and conventional modes of reckoning and thinking in terms of time. Now the integral concept of time cannot be fully explored by such tests which examine concepts having their origins in earlier years and seek to establish the age at which such concepts appear to become fully mastered. The purpose of this present test is to see if an essentially new, comprehensive concept appears at a certain age : new in the sense that it is a truly abstract and generalised notion of objective time. The theorectical basis for this hypothesis has been given above. Given this hypothesis, it was necessary to choose tests which would depend as little as possible upon knowledge of verbal symbols and conventional modes of time reckoning, and which would do more than test the precision of the subjects' understanding of historical time. Ideally, such items would present problems involving time which would be insoluble unless a clear and complete concept of objective time was possessed by the subject. Actually, most of the questions in the present test will be found to fall short of this ideal, but, as we have already said, equal weighting has been given to types of question which contain specific kinds of impurity in an attempt to minimise the effects of elements unrelated to the time concept.

The type of question required seems to find its best expression in the question asked by Michaud (quoted by Fraisse, 1957) : " Do we get older when we put the clocks forward in Summer ? " Questions 4 and 5 asked this same question and went a step further by asking the reason for the answer. To deal with this kind of probem, the subject must have an autonomous central concept of time as something objective and quite independent of clocks, calendars and other concrete phenomena which we normally employ in our temporal behaviour. To answer, as many did, that the clocks are put back again in Winter, so cancelling out the alteration, betrays the fact that the time concept is not yet completely autonomous and that the subject is still dealing with time at the level of concrete operations.

The dimensions of this article do not permit a full discussion of all the items in the test of the integral concept but the complete test is included in the appendix. The coefficient of reliability was only 0·5, but if the number of items had been greater, without destroying the balance or introducing new complexities, this could probably have been raised.

The questions in the test of the historical concept fell into four categories: historical sequence, historical analogy completion, temporal absurdity and the logical reconstruction of an historical account from scrambled data. All except the fourth category have been used many times before (cf. Sturt, 1922 ; Pistor, 1939 ; Friedman, 1943-1945 ; Rawlins, 1956). The fourth category is really an extension of the temporal absurdity, except that it demands a much finer discrimination in the locating of the absurdities and also requires that the events which are out of the proper chronological order should be reordered correctly from the information included in the scrambled account in the form of temporal cues and indications of cause and effect. It is true that this part of the test places a premium upon the capacity for verbal logic to quite a high degree, but it still depends upon the capacity to think in terms of historical time.

The test of the historical concept can be said to have more than face validity in that most of the items have been used by many previous investigators who have agreed in their definition of the historical concept. Moreover, the concept of historical time is generally recognised to exist whereas, in the case of the integral concept, the test was designed upon the hypothesis that such a concept might exist. The test of the historical concept had a coefficient of reliability of 0·88.

The total population tested consisted of approximately 330 children, some few children being absent from school for one test and present for the other. This population was divided into four groups according to age. The number of boys and girls in the various groups was roughly equal. Intelligence quotients were obtained from school records and date from the time of entry into the school. A description of the population is given in Table 1.

TABLE 1

THE POPULATION TESTED.

Group	Average Age	N	Average I.Q.	S.D.I.Q.
A	12·6	80	118·4	7·71
B	13·6	106	119·0	8·95
C	14·6	105	113·0	8·71
D	15·6	40	119·6	9·59

It will be noted that the average I.Q. is much the same for each group, except group C, where it was found necessary to include an extra form, lower in the academic scale, to obtain a comparably large N. In point of fact, since the correlation between time concept tests and I.Q. has been shown by other writers to be quite high (cf. Friedman, 1943, ; Rawlins, 1956) any consideration of the lower average I.Q. of group C could only be favourable to the kind of inference drawn and interpretation made below. The standard deviations of I.Q. show that variability within the groups was similar.

Recent work on intelligence has done much to undermine the older view of intelligence as fixed and predetermined. Piaget's work, in particular, has implied that intelligence consists of the possession of certain schemata, operations and concepts which determine problem solving, and that a child's level of intelligence reflects the speed with which he acquires these schemata and concepts. Indeed, McV. Hunt suggests that: " If Piaget's successive structures prove to have a fixed order, a very useful ordinal scale of intelligence would result. The times between successive landmarks of transition could then provide an inverse index to the capacity of various kinds of child environment interaction to foster intellectual development " (Hunt, 1961). It was with these considerations in mind that it was decided to select a population of above average intelligence, for if what we are seeking to define is the adult concept of time and its ' threshold ' age, then a more intelligent population would be most likely to approximate to the adult in its time concepts.

The socio-economic background of the population was homogenous: nearly all came from working-class homes, and in view of the findings of Friedman in this respect, it seemed this was a factor which could safely be ignored.

The tests were presented in school time, under normal classroom conditions in lieu of normal lessons. Invigilation was carried out by members of the staff of the school, in more than half of the cases by the author. All the marking was done by the author. Ample time was allowed for the tests so that no child failed to complete Test I and only eleven failed to complete Test II. (Test I: integral concept.)

III.—STATEMENT OF RESULTS.

From the group scores on Test I means and standard deviations were calculated and are shown in Table 2.

TABLE 2

GROUP SCORES—TEST 1.

Group	N	Score	Mean	S.D.	General Mean	General S.D.
A	81	1148	14·17	3·230	15·795	3·183
B	106	1567	14·78	3·018		
C	105	1756	16·72	3·586		
D	40	700	17·50	2·450		

The significant features emerging from Table 2 are the marked improvement between groups B, average age 13.6 years and group C, average age 14·6 years, and the fact that group C has the greatest degree of variability, which probably results from the wider range of I.Q. within that group which was commented upon earlier.

The data in Table 2 were next subjected to analysis of variance procedures, the results of which are given in Table 3.

TABLE 3

SIGNIFICANCE OF DIFFERENCE BETWEEN MEANS—TEST I.

Groups	Mean Difference	SE_D	t	Significance
A and B	0·61	·47	1·30	Not significant
B and C	1·94	·44	4·42	Significant at ·01 level
C and D	0·78	·59	1·32	Not significant
A and C	2·55	·47	5·43	Significant at ·01 level
A and D	3·33	·62	5·41	Significant at ·01 level
B and D	2·73	·59	4·60	Significant at ·01 level

The important feature here is that the difference in the means of groups B and C, which have a difference in average age of exactly one year, is significant at the ·01 level. It should also be noted that there is a similarly significant improvement in scores throughout the age range represented by this population taking it as a whole.

From the group scores on Test II, the test of the historical concept, means and standard deviations were calculated and are shown in Table 4.

TABLE 4

GROUP SCORES—TEST II.

Group	N	Scores	Means	S.D.	General Mean	General S.D.
A	80	1754	21·93	4·754	23·150	4·081
B	103	2400	23·30	3·033		
C	95	2242	23·60	4·482		
D	34	802	23·59	3·975		

The improvement shown between group A, average age 12·6 years and group B, average age 13·6 years, is the important element in the data derived from the test of the historical concept. Since the means otherwise were more or less identical, it only remained to calculate the significance of the difference between means A and B. The difference of 1·37 proved significant at the ·05 level.

The population was divided into boy and girl groups and subjected to the same procedures as were applied to the age groups. Data thus derived are shown in Table 5.

TABLE 5

DIFFERENCES BETWEEN SCORES OF BOYS AND GIRLS ON TESTS I and II.

	Boys			Girls			
Tests	N	Means	S.D.	N	Means	S.D.	Differences
I	164	16·19	3·365	168	14·97	3·343	1·22
II	158	23·47	4·101	154	22·66	4·111	0·81

The mean difference of 1·22 between boys and girls on Test I was found to be significant at the ·01 level.

The correlation coefficient for the two time concept tests was found to be 0·42.

IV.—DISCUSSION.

It was suggested above that previous writers have tended to equate the concept of time with the concept of historical time. The present test results would seem to cast doubts on that position. The test of the integral concept revealed a significant improvement between ages 13·6 years and 14·6 years, a year later than a similar, but less significant improvement in respect of historical time. The correlation coefficient between the two tests is as high as one would expect considering that they both deal with aspects of time, but not high enough to suggest that they are identical in scope. The inference to be drawn, therefore, is that the concept of historical time does not represent the ultimate in the development of time concepts as they feature in the daily psychological life of individuals in our society. For at least a whole year after age 13·6 in the population tested, notions of time continue to elaborate, the ability to generalise about time to increase and the flexibility of responses to stimuli involving time to expand at quite a rapid rate. Furthermore, whereas the means of the three older groups on Test II form a plateau after the initial improvement between groups A and B, suggesting that the elaboration of the concept of historical time has reached a peak, the means on Test I show a steady improvement throughout, implying that, although the main advance may take place between 13·6 years and 14·6 years, elaboration of the integral concept continues. Of course, the population tested was of fairly high average I.Q. and it is possible, since moderately high correlations have been found between tests of time concepts and I.Q., that the improvement in scores on Test I would perhaps appear later in a less selective population. Indeed, the possibility exists that individuals of much lower I.Q. might never master the concept at all. We have already pointed out that Paiget's work undermines the view of intelligence as a fixed dimension of a person and suggests that it relates to the age at which persons acquire certain mental structures and concepts, and to the kind of environment to which persons are subjected. The present findings in respect of time concepts may well have some bearing upon such an hypothesis, with regard to intelligence. If Piaget is right and intelligence is commensurate with the elaboration of concepts and schemata in a continuous process of interaction with the environment, " Of what," McV. Hunt asks, " do the individual differences in adult intelligence consist when formal operations are part of the equipment of all but the feeble-minded ? " (Hunt, 1961, p. 257). Hunt subsequently answers his own question with the proposition that intellectual development is, in fact, an open-ended process so that the differences at the adult level are essentially the same as those at the child level: the reflection of the stage of development of his concepts and schemata that the individual has reached (Hunt, 1961, pp. 309-310). Hunt observes elsewhere that the factor analysts have not revealed the existence of Piagetian structures and suggests that this may be so because their studies have not ranged across age groups and, therefore, that they were dealing with subjects who all shared the same basic structures (Hunt, 1961, p. 257). Placed in such a context, the present findings could well be interpreted to show that the historical concept is a schema absorbed about a year later into a higher schema, the integral concept of time, and that, whatever the I.Q. or age of the population tested, the sequence would remain the same. Finally, if the nature of intelligence is as Piaget's work implies, then the exploration of the more difficult concepts such as time could well play a part in the delimitation of intelligence at age levels beyond the 11 plus, which has hitherto been regarded as a watershed. Such speculations underline the need to extend this present type of enquiry to populations of lower I.Q., and indeed, of greater age. That the idea of an integral concept of time can be fitted into this context is perhaps an additional degree of validation for the test in question. Enlargement of Test I and the inclusion of a greater variety of items would, perhaps, produce more definitive results and would certainly be called for in any further investigation of the concept of time at this level. Nevertheless, the results of the present enquiry seem to justify the drawing of a clear distinction between the concept of historical time and the integral concept of time as stages in a process of conceptual development.

Previous writers have found no significant difference between the performances of boys and girls in temporal tests with the exception of Friedman (1945) who found that boys did better than girls on his historical time line sub-test. On the present test of the integral concept of time, the boys did considerably better than the girls. Now it is generally accepted that boys are usually better than girls at mathematics and some scientific subjects with a large mathematical content (cf. King, 1963). If this be so, then it is feasible that, since mathematics must depend upon the grasp of the numerical and spatial qualities of the external world and the ability to form abstractions of them, and since time itself is a dimension of that world and probably the most difficult to abstract, then boys could be reasonably be expected to show a readier understanding of time concepts. The fact that no previous writers have found such a difference only supports the present contention that previous tests have not explored the concept fully.

V.—REFERENCES.

AMES, L. (1946). The development of the sense of time in the young child. *J. Gen. Psychol.*, 68, 97-125.

BERGLER and ROHEIM (1946). The psychology of time perception. *Psychoanalytical Quart.*, 15, 190.

BRADLEY, N. C. (1947). The growth of knowledge of time in children of school age. *Brit. J. Psychol.*, 38, 67-78.

BUCK, J. N. (1946). The time appreciation test. *J. Applied Psychol.*, 30, 388-399.

CLAUSEN, J. (1950). An evaluation of experimental methods of time judgment. *J. Exp. Psychol.*, 40, 756.

FRAISSE, P. (1957). *Psycholgie du Temps*. University of France.

FRIEDMAN, K. C. (1943-4). Time concepts of elementary school children. *Elementary School J.*, 44, 337-352.

FRIEDMAN, K. C. (1944). Time concepts of junior and senior high school pupils and adults. *School Review*, 52, 233-238.

FRIEDMAN and MARTI (1945). A time comprehension test. *J. Educ. Res.*, 39, 62-68.
JAHODA, G. (1963). The child's concepts of time and history. *The Educ. Rev.*, 15.
JOHNSON, Ed. E. (1964) Time concepts as related to sex, intelligence and academic performance. *J. Educ. Res.*, 57, 377, 380.
KING, W. H. (1963) Science concepts. *Brit. J. Psychol.*, Nov.
LOVELL, K. (1960). *The Development of Mathematics and Science Concepts in Children.*
PIAGET, J. (1946). *Le development de la notion de temps chez l'enfant.* University of France.
PISTOR, F (1939). Measuring time concepts of children. *J. Educ Res.*, 33, 293-300.
RAWLINS, P. (1956). *An Enquiry into the Development of the Concept of Time in Children at a Secondary Modern School.* Sheffield Institute of Educ.
SCHACHTER, SYMONDS and BERNSTEIN (1957). The development of the concept of time in school children. *J. Nervous and Mental Diseases*, 7, 121.
SMYTHE and GOLDSTONE (1947). The time sense. *Perceptual and Motor Skills*, 7, 49-59.
STURT, M. (1922). *Concepts of Time.*
WALLACE AND RABIN (1960). Temporal experience. *Psychol. Bull.*, 57, 213-236.

(*Manuscript received*, 30*th May*, 1966)

VI.—APPENDIX.

A TEST OF THE INTEGRAL CONCEPT OF TIME.

1.—A time machine is featured in a current T.V. serial. What is a time machine ?
(*a*) A machine which can take you into the past or future.
(*b*) A machine which measures time.
(*c*) A machine which can make time go fast or slow.
(*d*) A machine for making a bomb go off at a certain time.

2.—What is a leap year ?
(*a*) A year of 364 days.
(*b*) A year of 365 days.
(*c*) A year of 366 days.
(*d*) A year missed out.

3.—Why do we have leap years ?
(*a*) Because the world is slowing down.
(*b*) Because there are 365 days in the year.
(*c*) Because the sun goes round the earth a little faster each year.
(*d*) Because it takes the earth 365¼ days to complete its orbit round the Sun.

4.—Every year we put the clocks forward by one hour at the beginning of Summer. How does this affect us ?
(*a*) We get one hour older.
(*b*) We get one hour younger.
(*c*) We stay the same age.
(*d*) We get one day older.

5.—In your own words, give the reason for your answer to the previous question.

6.—An aeroplane flew round the world from East to West, at such a speed that it arrived back at its starting point after 24 hours. It left at noon on 3rd July. What was the date it arrived back ?
During the flight of this aeroplane the pilot noticed that :
(*a*) The sun rose twice.
(*b*) the sun set twice.
(*c*) It was noon all the time he was flying.
(*d*) His day seemed to last forty-eight hours.

7.—What is speed ?
   (*a*) Travelling fast.
   (*b*) Travelling slow.
   (*c*) The distance travelled divided by the time taken to travel it.
   (*d*) Miles per hour.

8.—What is the meaning of the following words ?
   century
   generation
   decade
   eternity
   ancient.

9.—What year would be 400 harvests ago ?

10.—How long ago is 500 B.C. ?

11.—Underline the earliest and the latest of the following dates :
   113 B.C., 61 B.C., 507 A.D., 1300 A.D., 2 B.C., 17 A.D.

12.—A stalactite is formed by the dripping of water containing limestone, usually from the roof of an underground cave. In time, this dripping causes a kind of stone icicle to appear. How long do you think this process might take ?
   (*a*) A day.
   (*b*) A year.
   (*c*) 100 years.
   (*d*) A week.
   (*e*) 1000 years.

13.—A car travels at 40 m.p.h. and a train at 60 m.p.h. Both arrive at the same place at 11 o'clock, after travelling 160 miles. The car started at .... The train started at ....

14.—It was once said that if the people of China marched ten abreast across a bridge, the procession would never stop. (*a*) Could this possibly be true ? (*b*) Give your reasons for your answer below.

15.—In 1950, a man was twice as old as his son. His son is now 34. How old is the man ?

16.—Water is running out of a twenty gallon tank at two gallons per hour and a tap is letting water into the tank at one gallon per hour. How long will it take to empty the tank, presuming it was full to start with ?

17.—" If time stood still " is a commonly used phrase. If time were to stand still, what would be the result ?
   (*a*) You would be paralysed. Your pulse would stop, etc., but the world would go on around you just the same.
   (*b*) Everything in the world would stop moving.
   (*c*) Everything in the world would cease to exist.
   (*d*) All the clocks in the world would stop.

18.—The light of certain stars may be seen in the sky at night but the stars themselves no longer exist. Explain the reason for this.

RESEARCH NOTES

THE RELIABILITY AND VALIDITY OF THE GOODENOUGH-HARRIS DRAWING TEST

By WILLIAM YULE, LINDA LOCKYER AND AHILYA NOONE
(*University of London Institute of Education*)*

I.—Introduction.

As part of a large scale study of educational handicap among school children on the Isle of Wight, a group of 131 children was selected at random from the total population of approximately 2,200 children born between 1st September, 1953, and 31st August, 1955. The children were tested during November, 1964, when their average age was 10 years 1 month, and the group as a whole was found to be representative of the general population in terms of their socio-economic class as defined by paternal occupation.

Included in the battery of psychological and attainment tests administered to these children were a short version of the W.I.S.C. consisting of four subscales (Similarities, Vocabulary, Block Design, and Object Assembly) chosen on the basis of Maxwell's (1959) factor analysis, and the revised Goodenough-Harris Drawing Test which now contains 73 instead of 51 scorable items (Harris, 1963). The purpose of this study was to investigate the reliability of the new scoring system and to consider Harris's assertion that " . . . the correlation between an individual intelligence test result and the Goodenough-Harris Drawing Test score is quite substantial for children between the age of 5 and 10 . . . "

II.—The present study.

In accordance with Harris's instructions for mastering the scoring system, three raters independently studied his scoring criteria and sample drawings. Then, each scored the twelve test drawings provided in the test manual. Rank-order correlations of +0·99 were obtained between the point scores of each of the raters and those assigned by Harris. A one-way analysis of variance revealed no significant differences between the means of the three raters and Harris. It was, therefore, felt that the three raters had understood Harris's criteria for scoring all the points sufficiently well to justify a larger study of the inter-reliability of this scoring system and of the validity of the test as a measure of children's intelligence.

III.—Results.

(i) *Reliability*.

Each of the raters scored all 131 drawings independently. The total point scores were intercorrelated separately by age-group and sex, as can be seen in Table 1.

TABLE 1

Inter-Rater Reliability.

Group	N	Mean Age	Rater 1 *v* Rater 2	Rater 1 *v* Rater 3	Rater 2 *v* Rater 3
9½-year boys	33	9y 7m	0·92	0·84	0·87
9½-year girls	33	9y 8m	0·91	0·84	0·86
10½-year boys	31	10y 7m	0·89	0·93	0·88
10½-year girls	34	10y 7m	0·89	0·85	0·83
Total sample	131	10y 1m.	0·91	0·87	0·86

---

* This research was supported by grants from the Department of Education and Science and the American Association for the Aid of Crippled Children.

The inter-rater reliabilities range from +0·83 to +0·93 in the sub-groups, and from +0·86 to +0·91 on the total sample. These values are slightly lower than those reported by Harris.

Table 2 shows the results of an analysis of variance of the total point scores given to each drawing by the three raters. It can be seen that the null hypothesis, that there is no difference betewen the mean scores awarded to the total group by each of the three raters, has to be abandoned. Taken together, the results shown in Tables 1 and 2 indicate that whereas there is a fairly high agreement among the judges on their *ordering* of the different drawings, the absolute scores awarded differ significantly. This being so, it would be unwise to use one judge's score to estimate the intelligence of any individual child.

TABLE 2

ONE-WAY ANALYSIS OF VARIANCE OF DIFFERENCE BETWEEN MEAN POINT SCORES BETWEEN THREE INDEPENDENT RATERS.

Source	ss	df	v	F	P
Between groups	759·4	2	379·7	64·91	<0·001
Within groups	22,810·2	390	58·5		
Total	23,569·6	392			

(ii) *Validity.*

The sum of the four scale scores is taken as the measure of intelligence since age is controlled in converting raw scores to scale scores. Table 3 shows that there are large differences in the size of correlations between Draw-a-Man I.Qs. and W.I.S.C. I.Qs., both across age groups and across raters. Since this variation could well be due to sample fluctuations, only the overall correlations between the two estimates of I.Q. will be considered. Although these are statistically highly significant it is obvious that they are far too low to warrant the acceptance of the revised Goodenough-Harris Drawing Test as a test of general intelligence with children aged 9 to 11 years.

TABLE 3

PRODUCT-MOMENT CORRELATIONS BETWEEN DRAW-A-MAN I.QS. AND TOTAL. SCALE SCORES ON SHORT WISC.

	9½-year boys	9½-year girls	10½-year boys	10½-year girls	Total
Rater 1	0·21	0·43†	0·58‡	0·45†	0·41‡
Rater 2	0·13	0·34*	0·60‡	0·37*	0·34‡
Rater 3	0·14	0·21	0·66‡	0·38*	0·33‡

*—Significant at 5 per cent. level.
†—Significant at 1 per cent. level.
‡—Significant at 0·1 per cent. level.

IV.—REFERENCES.

HARRIS, D. B. (1963). *Children's Drawings as Measures of Intellectual Maturity.* New York: Harcourt, Brace and World.

MAXWELL, A. E. (1959). A factor analysis of the Wechsler Intelligence Scale for Children. *Brit. J. Educ. Psychol.*, 29, 237-241.

(*Manuscript received 17th March,* 1966)

REGRESSION ANALYSIS PROGRAMMES FOR THE IDENTIFICATION OF IMPORTANT VARIABLES IN EDUCATIONAL AND PYCHOLOGICAL EXPERIMENTS

By P. WILLIAMS
(*Faculty of Education, University College of Swansea*)

Probably the most important step in educational and psychological research is taken when the experiment is designed. In planning the design, one of the most important considerations is knowledge of those variables which are likely to play a major role and which will need controlling, through matching or randomising or other procedures. Usually, a search of the literature is helpful in reaching decisions about the relative importance of variables : sometimes it is also necessary to run a pilot experiment to investigate the field of research before the main experiment is designed. The purpose of this note is to draw attention to programmes of ' step wise regression ' (S.W.R.) which are of particular interest in this connection.

In a recent investigation concerned with reading development among educationally subnormal children, it was decided to run a pilot experiment to identify the most important variables for predicting reading improvement in this special group of children. Before running the pilot experiment it was felt wise to test a number of techniques of statistical analysis, including the step-wise regression procedure. Accordingly, a very short learning experiment (for the purposes of data collection) was carried out, in which four classes of fourth grade children in three communities in the Mid-West of the U.S.A. were taught Welsh orthography for one daily period of twenty minutes over four days. (This task was chosen since (i) it was expected that none of the children would have had any previous knowledge of it, enabling identical material to be presented to all the children ; (ii) learning could easily be measured by testing recall of the words taught over the four lessons and (iii) it offered the opportunity to introduce more interest into the lesson plans than would, say, an abstract learning task which might otherwise have been devised.)

Measurements were made on twelve predictor variables* and the complete data for seventy-nine children placed in an S.W.R. programme.

The programme used prints out the mean, variance and standard deviation of each variable, the covariance and correlation matrices. A regression equation is then computed for the criterion from the variable which predicts the greatest proportion of the criterion variance. Then the variable which predicts the greatest proportion of the remaining variance is introduced and a second regression equation computed. In turn, other variables are added singly to the regression equation until the last variable added makes no significant contribution to the variance being predicted.

In this way a series of regression equations are available for study. The programme prints out at each stage the multiple correlation coefficient (R) between predictor variables and the criterion, the constant term of the regression equation, the beta weights and the total and predicted sums of squares.

The series of regression equations which are obtained and the changes which occur as variables are added provide useful information. In the experiment mentioned above, the initial matrix of correlations is given below. (Variable 13 is the criterion variable, Welsh orthography score).

Next, we can look at the extent to which the addition of different variables improves the prediction of our criterion variable. This is shown in table 2.

DISCUSSION.

(1) The S.W.R. programme used variables 9 and 1 only since it was at this point that the addition of the next variable produced no significant increment in variance. The details of the addition of the remaining ten variables are given in the table above for interest only, and the programme was modified for these details to be obtained.

---

* This experiment will be reported more fully in Williams, Ross and Sadek.

TABLE 1

Matrix of Intercorrelations (Decimal Point Omitted, Negatives in Italic Figs.).

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1		13	35	*53*	03	*03*	22	26	*26*	*22*	*26*	*23*	*51*
2			06	03	*02*	26	*03*	16	*15*	*01*	*20*	*05*	*04*
3				*97*	*17*	01	12	05	10	28	08	*03*	*12*
4					15	03	*16*	*09*	05	*20*	*04*	*09*	22
5						*03*	09	*03*	*07*	*27*	18	*21*	*09*
6							13	*00*	*48*	*09*	*08*	01	*15*
7								10	23	*06*	06	15	*14*
8									*31*	*16*	*25*	*01*	*25*
9										45	44	33	55
10											74	55	39
11												53	38
12													36

The last column in this table represents the correlation of the different variables with the criterion variable.

TABLE 2

Sums of Squares (to First Place of Decimals) Associated with the Addition of Each Variable to the Prediction Equation: Multiple Correlation Coefficients.

Variable	Additional Sum of Squares Predicted	R
9	263·6	+0·55
1	125·3	+0·67
10	12·2	+0·68
4	5·6	+0·68
3	3·3	+0·69
7	2·9	+0·69
6	0·7	+0·69
5	0·5	+0·69
8	0·2	+0·69
11	0·2	+0·69
12	—	+0·69
2	—	+0·69

Total Sum of Squares Predicted=414·5.
Total predictable Sum of Squares=867·4.

(2) The regression equation given at this point was:

$$\text{Variable } 13 = +0 \cdot 11 \text{ (Variable 9) } -26 \cdot 1 \text{ (Variable 1)} + 778$$

More complicated regression equations using more predictor variables can be calculated, but they are hardly more effective than this one.

Note how the addition of other variables makes little difference to the multiple correation coefficient notwithstanding the fairly high correlation of some of these variables with the criterion variable. Thus, although variable 11 correlates quite well (+0·38) with the criterion variable, its introduction into the prediction (regression) equation (in this particular sequence) results in a quite insignificant increment in the sums of squares predicted and in R.

(3) If this were more than an exercise in data analysis we would be justified in attaching considerable importance to controlling variables (9) and (1) in any experiment designed, for example, to assess the relative value of different methods of teaching Welsh orthography. But we would also want to reanalyse the data in other ways (e.g., by one or more types of factor analysis) in order to obtain more insight into the interrelationships, and to use more variables in the hope of predicting a greater proportion of the criterion variance.

(4) Various S.W.R. programmes exist which will differ, for example, in the nature and cut-off point of the test of significance employed. Details such as this will depend on the task for which the programme was first written. The programme used in this example was written in Fortran 60, and is similar to that described by Efroymson in Ralston and Wilf (1960).

(5) Other types of regression analysis programmes offer more sophisticated ways of looking at sets of data, e.g., through computing the proportion of the criterion variance predicted by each one of all possible combinations of predictor variables. Details of programmes available can be obtained from the users' organisations.

(6) Although, as with all statistical procedures, there are some snags in the use of S.W.R. (see, for example, Cooley and Lohnes (1962)), nevertheless, it is a procedure which has considerable application in areas of research in educational psychology where multivariate analysis is needed.

REFERENCES.

[1] COOLEY, W. W., and LOHNES, P. R. (1962). *Multivariate Procedures for the Behavioural Sciences.* Wiley, p. 35.

[2] RALSTON, A., and WILF, H. S. (1960). *Mathematical Models for Digital Computers.* Wiley.

[3] WILLIAMS, P., ROSS, B., and SADEK, F. (in preparation). The development of a research programme for materials evaluation.

(*Manuscript received, 17th March,* 1966)

THE MINNESOTA TESTS OF CREATIVE THINKING—A NOTE ON SCORER RELIABILITY IN FOLLOW-UP STUDIES WITH ENGLISH PRIMARY SCHOOL CHILDREN

By R. J. GOLDMAN
(Didsbury College of Education)

AND D. F. CLARKE
(Newland Park College of Education)

SUMMARY. This study is part of a series of pilot investigations into various aspects of creative thinking with small groups of primary school children in this country using mainly Minnesota Creative Test material. The object of this enquiry was to discover more about the reliability of subjective interpretations of written material by a single and independent scorers. The results of this study indicate that a high level of reliability in judgments of creative writing obtains for the individual and between independent assessors, so long as the criteria for scoring are clearly understood.

INTRODUCTION.

The Minnesota Tests of Creative Thinking, devised by E. P. Torrance (1962), include an Imaginative Stories Task and a battery of verbal and non-verbal tests. The tests have been described and evaluated by Goldman (1964). They are designed to reveal various aspects of creative thinking in the child, and a comprehensive scheme of marking has been drawn up. The purpose of this enquiry was to discover if the interpretations of individual assessors of children's imaginative stories are reliable between two marking sessions and if independent assessors compare closely in their judgments of imaginative story writing and verbal and non-verbal tasks by children, when clear criteria are laid down.

PROCEDURE.

The Imaginative Stories task of the Minnesota Tests of Creative Thinking was administered by the headmaster to thirteen boys and ten girls aged between 10 and 11 years in a primary school in a mixed urban-rural area. Each child was asked to write the most interesting and exciting story he or she could think of about one of the following topics: (1) " The Dog that doesn't bark "; (2) " The Man who cries "; (3) " The Woman who can, but won't talk "; (4) " The Cat that doesn't scratch "; (5) " Your Teacher stopped teaching "; (6) " The Doctor who became a Carpenter "; (7) " The Cock that doesn't crow "; (8) " The Horse that won't run "; (9) " The Duck that doesn't quack "; (10) " The Lion that doesn't roar ". The time limit was 20 minutes. The Revised Originality Scale for evluating Creative Writing was used, based upon the six rating scales by Yamamoto (1961). Grades 0–2 were given to each of the aspects listed below. Where the writing showed none of the qualities under consideration, the value O was given. Clear manifestations were given 2 and those which were considered to be poor, weak or obvious and banal were awarded 1. The following guide has been taken from E. P. Torrance, R. G. Peterson and D. J. Davis (1963):

(1) Unusual titles.
(2) Picturesque speech.
(3) Vivid.
(4) Flavor.
(5) Personal element.
(6) Original solution or ending.
(7) Original setting or plot.
(8) Humor.
(9) Invented words and names.
(10) Individuality of style.
(11) Becomingness.
(12) Imagination.
(13) Emotions or feelings.
(14) An abstractive element consisting of finding the essence.
(15) Immediacy of experience.

(16) Curiosity.
(17) Reservoir of experiential data.
(18) Perceptive sensitivity.
(19) Flexibility or versatility.
(20) Symbolism.
(21) Coherent unity.
(22) An expressive—communicative element.
(23) Unique punctuation.

The material was marked by the writer, the results sealed, and after one week, the stories were assessed again. The scripts were then sent to the headmaster of the school where he marked the stories, sealed the results, and waited a week before marking the stories again. The scores were then correlated as follows :

(*a*) 1st evaluator—1st and 2nd assessments.
(*b*) 2nd evaluator—1st and 2nd assessments.
(*c*) 1st evaluator's average scores for each script with the 2nd evaluator's average scores for each script.

Boys' and girls' scores were correlated separately. The individual scores, transforms for Z, standard errors of Z and standard errors for comparing any two Zs, have been omitted from this paper in the interests of brevity.

RESULTS.

With the exception of the first evaluator's two assessments of the girls' stories, the high correlations support Burgart's (1961) view that independent judges who have previously discussed criteria for marking arrive at very similar interpretations. Although Burgart's findings were in respect of interpretations of paintings, a similar degree of reliability would seem to operate with judgments of creative story writing. The " Revised Originality Scale for Evaluating Creative Writings," therefore, appears to furnish a useful guide for the interpretation and scoring of imaginative stories. The lower scorer reliability in respect of the girls' marks (1st evaluator) suggest that caution should be observed when discussing any scores, even though high reliabilities generally obtain.

TABLE 1

SCORER RELIABILITIES.

	Boys. n=13	Girls. n=10
1st Evaluator—1st and 2nd scores ..	·921	·484
2nd Evaluator—1st and 2md scores ..	·947	·949
1st and 2nd Evaluators (average scores)	·861	·721

W.R. Palmer's results (1964) also show generally high degrees of reliability in regard to independent scoring, ranging from 97 per cent. to 99 per cent. agreement in fluency and flexibility (verbal and non-verbal) down to more modest agreements in originality (verbal 56 per cent. and non-verbal 67 per cent.). The factor showing widest variance was elaboration (non-verbal 83 per cent. and verbal 49 per cent.). Small differences in scores would, of course, be exaggerated when converted to percentages.

Taylor (1965) in her study, observed high scorer reliabilities of between ·93 and 1·00 between four markers. Again Elaboration seems to have been the most difficult factor upon which to secure agreement, although the present writers have noted high correlations of ·876 (boys) and ·952 (girls) for elaboration in another study which involved the Tin Cans Test. In this, two evaluations by the same marker were correlated.

CONCLUSIONS.

(1) In general, there seems to be a high level of scorer reliability when the evaluator clearly understands the criteria for scoring.

(2) There appears to be a reasonably high level of reliability in the assessments between independent scorers when criteria are discussed and understood.

(3) High levels of reliability in scoring between independent markers are also generally observed by W. R. Palmer, B. Taylor, and K. Yamamoto in their studies of creative thinking.

REFERENCES.

BURGART, H. J. (1961). An analysis of the relationship between acquired skill and confidence; general creativity and aesthetic performance at the higher education level; the relationship of art experience to personality, general creativity and aesthetic performances. The Pennsylvania State University. Reprint from *Dissertation Abstracts*, Vo. XXII, No. 71962.

GOLDMAN, R. J. (1964). The Minnesota tests of creative thinking. *Educ. Res.*, VII, 1, 3-14.

PALMER, W. R. (1965). *An Investigation into the Creative Abilities of a Group of* 10–11-*year-old Children in a New Town.* Maria Grey College of Education, July, 1965.

TAYLOR, B. (1965). *A Study of the Creativity and Intelligence of Children in the Fourth Year at a Primary School.* A special study submitted in the Department of Education, University of Reading.

TORRANCE, E. P. (1962a). *Guiding Creative Talent.* Prentice Hall, Inc., Englewood Cliffs, N.J.

TORRANCE, E. P. (1962b). *Abbreviated Form VII Manual.* Bureau of Educational Research, University of Minnesota Publication.

TORRANCE, E. P., PETERSON, R. G., and DAVIS, D. J. (1963). *Revised Originality Scale for Evaluating Creative Writing.* Bureau of Educational Research, University of Minnesota.

YAMAMOTO, K. (1961). *Manual for Everlasting Imaginative Stories.* Bureau of Educational Research, University of Minnesota.

(*Manuscript received 24th March*, 1966)

THE PROGRESSIVE MATRICES (1938) AND THE DOMINOES (D48) TESTS: A CROSS-CULTURAL STUDY

BY A. ABI RAFI

(*Aley, Lebanon*)

The Progressive Matrices (1938) and the Dominoes (D48) tests were used within a battery in the selection of candidates for entry to a technical college. The present data cover four such competitions and comprise 1,167 Lebanese men aged 17 to 25. The PM was used as a speed test with a time limit of twenty minutes. The time for the D48 was twenty-five minutes.

Results* show that the rank order of difficulty of the PM maintain item position within the sets as well as between sets. On the D48 the middle range of items were found more difficult than their item position indicated. Items 17, 18, 23, 24, 25 and 26 are ranked 25th, 24th, 29th, 30th, 35th, and 38th, respectively, in the ranking based on order of difficulty.

A comparison was made between the present ranking order on D48 and that of Gough and Domino (1963). The coefficient of ranked correlation was ·948. Considering that Gough and Domino obtained rank order correlation of +·83 and +·89 between their sample and data of two European studies, it appears that there is a fairly strong evidence against the present order of difficulty and item position of this test.

---

* Extended data can be obtained from the author.

Another sample was tested which comprised 170 males comparable with the larger sample both in age range and educational level Here the PM was used as a power test following Raven's (1956) instructions. The time for the D48 remained unchanged It was found that the items present the same level of difficulty for both samples. The coefficient of ranked correlation was ·946 for the PM and ·971 for the D.48.

In comparing the two tests with each other, the product-moment coefficient of correlation for the larger sample was ·580 with Mean 38·10, S.D. 8·85 for the PM, and Mean 26·04, S D. 5·80 for the D48. For the smaller sample, the coefficient of correlation was ·595 with Mean 42·00, S.D. 8·91 for the PM, and Mean 25·73, S.D. 5·97 for the D48. It is of interest to note that the PM mean and standard deviation of the smaller sample is similar to the mean and standard deviation of Raven's standardization population of 3,665 subjects for whom the respective figures were 42·74 ± 9·77 (Eysenck, 1947, p. 114).

The table compares the present results with those of other studies. Most of the figures are taken from Gough and Domino, Table 2, which summarized the results of European studies. Only those results that include male subjects of comparable age and educational level with the present samples, are included.

TABLE OF

ILLUSTRATIVE NORMS ON THE D48 TEST.

Samples	N	N	SD
1. From Ferracuti and Rizzo (1959) Italy			
(*a*) University students, males	100	26·70	5·60
2. From Maury (1954) France			
(*c*) Students, deuxieme classe (11th grade)			
(1) Males	60	27·70	4·68
(*d*) Students, Baccalaureat (12th grade)			
(1) Males	70	30·70	5·04
3. From Pasquasy and Doutrepont (1956) France			
(*a*) Young men ages 20-25	522	19·78	8·61
4. From Pasquasy and Doutrepont (1956) Belgium and Switzerland			
(*g*) Students, both sexes, age 16 · 6 and over	87	25·57	6·72
5. From Pichot *et al.* (1953) France (Baccalaureat complet)			
Males, age 20-24	98	30·13	6·00
6. From Domino (1964) U.S.A.			
College males	94	28·06	5·44
7. Present study, Lebanon			
(*a*) College grade, males	1167	26·04	5·80
(*b*) College grade, males	170	25·73	5·97

The correspondence between the average scores of studies 1, 2c and 3 with those of the present investigation further support the suggestion advanced by Gough and Domino that " . . . average scores for subjects of similar age and educational level in different countries will be about the same " (p. 346).

The principal implication of this investigation is that both tests yield results which might compare well cross-culturally.

REFERENCES.

EYSENCK, H. J. (1947). *Dimensions of Personality*. London: Kegan Paul, Trench, Trubner and Co.

DOMINO, GEORGE (1964). Comparison of the D48, Cattell Culture Fair, and Army Beta tests in a sample of college males. *J. Consult. Psychol.*, 28, 468-469.

GOUGH, H. G., and DOMINO, G. (1963). The D48 Test as a measure of general ability among grade school children. *J. Consult. Psychol.*, 27, 344-349.

PICHOT, P., *et al.* (1953). Quelque consideration a propos des tests D48 et Vocabulaire. *Rev. Psychol., Appl.*, 3, 395-405.

RAVEN, J. C. (1956). *Guide to Using Progressive Matrices* (1938). London: K. H. Lewis.

(*Manuscript received 11th July, 1966*)

CO-EDUCATION AS A POSSIBLE INFLUENCE ON THE DESIRE FOR MARRIAGE

BY R. R. DALE
(*University College of Swansea*)

SUMMARY. It was hypothesized that co-educational, as opposed to single-sex, schooling might have a slight effect on the desire for marriage. Estimates were obtained from student teachers in England and Wales. The differences, though interesting, were not statistically significant. The influence of social class and of siblings on the estimates was also examined.

I.—INTRODUCTION.

A questionnaire was administered to 2,000 student teachers in twenty-one institutions in England and Wales. One question was" Do you wish to marry, provided you can find a suitable partner ? " Estimates were requested on a seven-point scale ranging from " Very much /Much /A little /Undecided " to " Very much opposed." The estimates were then weighted seven for " Very much " to one for " Very much opposed." The mean scores of ex-pupils of single-sex grammar schools were compared with those from mixed grammar schools. Replies were anonymous and free comments were requested where opposition to marriage was signified. Married students were excluded.

II.—WOMEN'S ESTIMATES.

The hypothesis was that social interaction between the sexes at school might slightly increase their desire for marriage, compared with pupils from single-sex schools. For the 972 women the difference was unexpectedly in the oppsoite direction, though not statistically significant (Chi-square was 5·832 for 6 degrees of freedom). 67·4 per cent. of the girls from girls' schools wished " Very much " to marry, and 61·5 per cent. from mixed schools, but most of this difference was reversed in the ' Much ' estimates. In both groups only about 2 per cent. gave any of the three ' Opposed ' estimates. The results of another question on ' Wish for children ' lend support by their close similarity. It may be that deprivation—or lack of contact with the opposite sex—slightly increases the desire for marriage and parenthood amongst women, but confirmatory evidence is needed. The effect is interesting but too small to be of practical significance.

As the social class distribution is different in single-sex compared with mixed grammar schools, part of the information was analysed according to this variable. No significant social class differences were found, nor did the variables have any appreciable effect on the main comparison except to reduce the difference very slightly. These data were then divided into ex-pupils ' With siblings ' and ' Without siblings,' for the parental skilled artisan class only. Ex-pupils of girls' and mixed schools were combined, as their means and distributions were so similar. The mean score of those ' with siblings ' (163) was 6·42, and ' without siblings ' (56) 6·50, but the difference was not statistically significant. Almost 6 per cent. of those with siblings were undecided or opposed, while none of those without siblings were opposed and only 3·6 per cent. were undecided. Perhaps this is again a slight hint that lack of contact with the opposite sex is a rather more powerful influence than is the actual experience of it—which may, after all, be tragic or unhappy in some cases.

Only seventeen comments are given for opposition to marriage. The general theme is sometimes a frank avowal of selfishness, but more often a mere wish to stay independent or to indulge in an alternative, such as travel. Fear of having children and dislike of the idea of sexual intercourse are both mentioned. Noteworthy are: " I'm not made for self-sacrifice—natural outcome of marriage for women." " Too limited—no scope for independent activity." " There are so many things I want to do that to be tied down to one person would not be good for me or him," and " I cannot see myself settling down and being happy with one person for the rest of my life."

III.—MEN'S ESTIMATES.

In contrast to the women, the men's results were in the direction of the hypothesis that co-educational schooling would make them slightly more favourable to marriage. The difference between the two groups was, however, very small and not statistically significant. Of the 304 men from mixed schools, 63·3 per cent. wished to marry ' very much,' while of the 479 men from boys' schools only 59·5 per cent. were in this category. A mere 0·6 per cent. of the mixed schools group were opposed to marriage, and 2·3 per cent. from the boys' schools. When the two distributions were divided into the favourable and the opposed, the difference between men from mixed and men from boys' schools gave a Chi-square of 3·07, which, for 1 degree of freedom, is somewhat short of statistical significance.

Analysis by father's social class produced a striking similarity between the classes within each type of school, but a persistent though small difference between the schools within each social class. These differences, taken individually, were not statistically significant, but support each other by their consistent pattern. Both here and in the women's section we must remember that respondents were all intending teachers, and therefore, moving towards the same social class, so that any differentiating force of social class would thereby be reduced. Of the 783 who participated, only thirteen opposed marriage. Two of these thirteen came from mixed schools, and wrote: " I feel I want to carry on my studies so that I am taxed fully. To have to devote time to a wife and family would hinder me " (A little opposed), and " I enjoy my own company and freedom. I don't want the responsibility." (Much opposed). Of the eleven men from boys' schools who opposed marriage (out of 479 who answered), three failed to give their reasons, though requested to do so. The chief reasons given are loss of freedom, both personal and financial. There was one objection to ' monogamous marriage,' and one man was ' appalled at the thought of such a relationship." Another was " Very much opposed " because he was too moody and because his parents' marriage had been an unhappy one.

IV.—COMMENT.

The slight tendency to greater variability of scores among the group from single-sex schools is interesting and in the case of the men approaches statistical significance. With the latter some opposition to marriage could be caused by a tendency to homo-sexuality among a small fraction of the pupils in boys' schools. Although there is a very slight indication of this in one comment, the evidence is inadequate for any definite conclusion. It was thought that the number of respondents from boys' schools who failed to reply might provide another pointer, when compared with the number of male non-respondents from mixed schools, on the hypothesis that homo-sexuals would refuse to answer a question on marriage. The difference was in the expected direction—2·1 per cent. from boys' schools, compared with 1·3 per cent. of the boys from mixed schools, but it is not statistically reliable,* nor in any event could one be at all sure that the difference was due to the homo-sexual factor. Complaints regarding homo-sexuality tend to come, however, from boarding rather than day schools, and the former represent only 5 per cent. of the present sample from boys' schools. Of the ten respondents from boys' schools who failed to give any estimate, only one was a boarder. The questionnaire was not, of course, designed to elicit information on this specific point; and even if homo-sexuality were an underlying cause of opposition it is unlikely that this reason would be stated.

---

* Chi-square = ·67 for one degree of freedom.

(*Manuscript received 4th August, 1966*)

THE USE OF VALENTINE'S REASONING TESTS IN AN INVESTIGATION INTO TRANSFER FROM MATHEMATICS TO REASONING

By J. K. BACKHOUSE
Department of Education, University of Oxford)

Summary. The writer wished to test the widespread belief that mathematics provides a training in reasoning. A favourable opportunity to do this is provided by the specialisation in sixth forms when many pupils drop mathematics and when there is the possibility of observing comparatively long-term effects.

I.—Method.

Valentine's *Reasoning Tests for Higher Levels of Intelligence*[1] was taken as a criterion of reasoning ability and a parallel test was constructed, sixty-eight first-year training college students working the final version. Split-half reliability coefficients were: 0·75 for Valentine's test and 0·73 for the parallel test; the correlation coefficient between the scores was 0·73. The parallel test has not been published.

A random sample of twenty maintained boys' grammar schools was drawn from counties near Oxford; fourteen of these schools were able to co-operate by administering Valentine's test during the first fortnight of the Autumn term, 1963, and the parallel test during the Summer term of 1965. Information was available about the boys' choice of subjects for their courses leading to Advanced Level of the G.C.E.

II.—Results.

Scores on tests. 613 boys sat the first test in 1963; their scores had mean 30·0 and standard deviation 13·7, the school means varying from 23·4 to 39·1.

513 boys took both tests and their mean gains in the fourteen schools varied from 5·5 to 17·6, the mean gain being 11·5. The long-term practice effect is not known (in the final evaluation and the penultimate trial of the parallel test, practice effects of six and nine were observed), but it would appear likely that some of the increase is due to the boys' greater maturity in 1965.

Gains in scores. Boys were classified as 'mathematicians' (M) or 'non-mathematicians' (NM) according as to whether their average number of periods of mathematics per week during term was five or more, or one or less, respectively; the others were ignored for this part of the investigation in order to highlight any possible transfer effects. The mean increase in score was slightly higher for the NM (11·5, $n=223$) than for the M (11·3, $n=226$), the same observation being true for ten out of the thirteen schools where a comparison was possible. This tendency is to be expected on the hypothesis of no transfer since there was a negative correlation (–0·290) between scores on the first test and gains between the two, and the mean of the M's on the first test was higher than the means of the NM's. When allowance was made for the higher mean scores of the M's on the first test, the corrected mean gain of the M's exceeded that of the N-'s by 1·6. Had the sample been random with respect to boys (instead of schools) the difference would not have been significant; this, together with an inspection of the data, suggested that an analysis of variance would not be worth-while.

A similarly inconclusive result was found for the corrected mean gain of those who had read about, or been taught, certain topics of ' modern mathematics ' (Boolean algebra, logic, Venn diagrams). Out of the 449 M's and NM's, the corrected mean of those with some acquaintance with any of these topics exceeded that of those without by 2·4.

Differences between arts and science sides. Defining ' science ' subjects as mathematics, physics, chemistry, biological subjects and technical drawing, and ' arts ' subjects as English, history, geography, economics, languages, music and art, it was found by an analysis of variance ($F=18{\cdot}3$, $P<0{\cdot}001$) that the mean score on Valentine's test of those taking two or more science subjects (32·4, $n=320$) was significantly higher than that of those taking two or more arts subjects (27·4, $n=293$). An examination of the high scores showed that: of those over 60, four were on the arts side and eleven on the science side; of those over 50, seventeen were arts and forty-four science; and of those over 40, forty-seven were arts and 102 science.

Differences between subjects. Thirteen of the schools had boys taking courses to Advanced Level of the G.C.E. in the subjects below; and analysis of variance showed highly significant differences between the mean scores (Valentine's test) taken by subject ($F=9{\cdot}97$, $P<0{\cdot}001$): mathematics 34·1, physics 32·6, chemistry 31·7, French and biological subjects 28·8, history 27·1, English literature 27·0, geography 26·9.

Other variables. No significant correlation was found between socio-economic class and score on the 1963 test, nor between age and test score; the age range was, of course, severely restricted by the sample.

III.—CONCLUSIONS.

A comparison of the inconclusive results on gains in scores and the highly significant differences between subjects suggests a firm conclusion: the average superiority of mathematicans (at least at school level) is better accounted for by their initial superiority at reasoning than by the transfer effects from mathematics to reasoning. The transfer effects from ' modern mathematics ' to Valentine's test deserve further investigation—the contact of some of the boys with the topics mentioned was very slight.

The order of mean scores taken by subject is striking. The trend of decreasing scores from mathematics through physical sciences and languages to other subjects, with biology coming in the region of languages, holds for fifteen subjects studied to Advanced Level of the G.C.E.; it may well correspond to the demands on logical reasoning made by those subjects.

Valentine[2] remarks that his reasoning test may be regarded as testing intelligence, in which case these results are in consonance with those of Hudson[3] on the distribution of intelligence between subject groups and with the findings of Watts[4] on the effects of a university training in mathematics on A. H. 5 scores.

ACKNOWLEDGEMENTS.—The writer wishes to record his thanks to: Messrs. Oliver and Boyd who, with the late Professor Valentine, allowed the use of copyright material in the preparation of the parallel test; the college of education principals, lecturers and students who helped with its trials; the headmasters, staff and boys of the participating schools for their co-operation over the tests, Mr. A. D. C. Peterson for making the research possible; Mr. A. Yates for his encouragement and advice; Mrs. P. M. B. Cuckney for assistance in marking tests; and Mr. G. B. Hill for data concerning the boys' socio-economic groups.

IV.—REFERENCES.

[1] VALENTINE, C W. (1954). *Reasoning Tests for Higher Levels of Intelligence.* Edinburgh : Oliver and Boyd.

[2] VALENTINE, C. W. (1961). The use of a new reasoning test for selection of university and training college students. *Brit. J. Educ. Psychol.*, 31, 227-231.

[3] HUDSON, L. (1963). The relation of psychological test scores to academic bias. *Brit. J. Educ. Psychol.*, 33, 120-131.

[4] WATTS, KATHLEEN P. (1954). The effect of a university training in mathematics on the scores in an intelligence test. *Brit. J. Educ. Psychol.*, 24, 32-37.

(*Manuscript received* 15*th August*, 1966)

COMMENT

NOTE ON THE ARTICLE "DIFFERENCES IN TEST SCORES AND IN THE GAINING OF SELECTIVE PLACES FOR SCOTTISH CHILDREN AND THOSE IN ENGLAND AND WALES," BY DOUGLAS, ROSS, MAXWELL AND WALKER.

By HUGH LYTTON
(*Institute of Education, University of Exeter*)

The authors of this very interesting article are surprisingly cautious in interpreting the actual test score differences between Scottish and English children, saying that they are " difficult to interpret because, of course, the tests used varied from year to year, not only in subject matter but also in type."

This caution is perhaps due to the fact that the authors kept their interpretation to the total T scores. Yet, if one examines individual test results some interesting conclusions emerge, conclusions which, in fact, are closely parallel to the findings of a previous investigation, based on much smaller numbers , by Vernon, O'Gorman and McLellan (1955).

In the National Survey of Health and Development analysis, the Scottish children were inferior to the English (here includes the Welsh) in non-verbal and verbal reasoning tests at all ages, except in verbal reasoning at 15. These differences are repeated *inside* classes, except in one case (middle class 8-year-old children—no difference). In the previous investigation, too, Scottish children did slightly less well in this type of test.

In spite of this inferiority, the Scottish children are clearly superior in mechanical reading and sentence completion (which depends on reading) at 8 years—as they were in the Vernon study. The mechanical reading test was repeated at 11 years and by then the Scottish superiority has disappeared. (No mechanical reading test at 11 in the earlier study but reading comprehension scores were equal.)

However, in arithmetic (which had not previously been tested) the Scottish 11-year-old children are superior in the N.S.H.D. investigation, as they were in the earlier study. At 15, a ' mathematics ' test replaced the arithmetic test. From a private communication from Miss J. M. Ross, it appears that this N.F.E.R. test was similar to Vernon's Graded Arithmetic—Mathematics Test, in which arithmetic is a major component skill. The Scottish arithmetic superiority then is seen to persist, and to grow into a Maths. superiority at 15, parallel to the findings by Vernon, *et al.*

It is generally accepted, as Vernon, *et al.* wrote, that " Scottish primary education tends to be more formal than English, involving the earlier introduction of, and more thorough drilling, in, the ' three R's '."

It seems to me one is justified in drawing the following conclusions from the N.S.H.D. results : the early formal teaching of reading conferred only very temporary advantages on Scottish children which disappeared by the age of 11, by which time incidental reading and reading for pleasure will provide as great a source of practice for most children as formal lessons at school.

The formal and systematic teaching of arithmetic bestowed advantages up to school leaving age, as in this area skill depends mainly on continued practice provided at school. It seems then that early formal teaching and thorough drilling provide advantages only where school-based practice continues to be the only source of reinforcement In skills less dependent on practice in school, the more formal teaching does not seem to confer any lasting benefits.

This interpretation is in line with the findings by Vernon, *et al.* (1955), although in that case the school-practised skill in which Scottish superiority was significant, was spelling. It is interesting to note that Gibson (1966) in a comparison between 11–12-year-old British (English, Scottish and Irish) and United States pupils found that the achievement area in which British pupils were consistently superior was ' arithmetic fundamentals.'

It would be invaluable to administer similar tests to the Survey subjects again at the age of 18 and/or 21 (perhaps this has been done). One could then see if the Scottish superiority in Mathematics survives several years of no practice and hence has given Scottish children a lasting advantage.

REFERENCES.

DOUGLAS, T., ROSS, J., MAXWELL, S., and WALKER, D. (1966). Differences in test score and in the gaining of selective places for Scottish children and those in England and Wales. *Brit. J. Educ. Psychol.*, 36, 150-157.

GIBSON, R. L., and FRAWZEN, W. L. (1966). Six to eleven: American and U.K. pupils compared. *New Society*, No. 200, 28th July, 1966.

VERNON, P. E., O'GORMAN, M. B., and MCLELLAN, A. (1955). A comparative study of educational attainments in England and Scotland. *Brit. J. Educ. Psychol.*, 25, 195-203.

(*Manuscript received* 31*st July*, 1966)

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES

THE ATTAINMENTS IN READING AND SPELLING OF CHILDREN WHO LEARNED TO READ THROUGH THE INITIAL TEACHING ALPHABET

By TERENCE D. SWALES

(*Abstract of thesis submitted for the degree of M.Ed., University of Manchester,* 1966)

The main aim of this investigation was to obtain evidence as to the efficiency of the Initial Teaching Alphabet (i.t.a.) on the three criteria of:

(*a*) Average achievements in reading comprehension.

(*b*) Average spelling attainments.

(*c*) Incidence of backwardness in reading.

METHOD.

A sample of children, taught by i.t.a., in six different schools, i.e., every available school in North Staffordshire using the new system and willing or able to co-operate, were tested at the end of their third school year. A control sample using t.o. drawn from another six schools was also tested. It was hoped to avoid as far as possible any Hawthorne effect. Neither group knew at the start that it was to be the subject of a future experiment. The method proposed for the main investigation was to compare the results of two group tests:

(1) N.F.E.R. Sentence Reading Test 1.

(2) Daniels and Diack Graded Spelling Test.

The groups were matched in respect of the following details: (i) number of children: 99; (ii) sex: 54 boys and 45 girls; (iii) age, in years: months: 7 : 4–8 : 1 (mean age 7 : 9); (iv) length of schooling: nine school terms; (v) social status: mean social class 3.5, according to father's occupations (Registrar General's Classification), school catchment area was also taken into account; and (iv) intelligence recorded by N.F.E.R. Picture Test 1: each group contained three matched subgroups of high, low and medium intelligence: the mean quotients were 120, 82 and 100, respectively.

Altogether 108 i.t.a. children and 124 t.o. children were tested. The matching was facilitated by selecting schools of similar size, type and catchment area. The i.t.a. group had begun reading by the Janet and John reading scheme (transliterated). Four of the t.o. schools had begun with 'Janet and John' and two by the Happy Venture Series. Both groups had used supplementary readers.

Analysis of variance was used to test crucial hypotheses. Three way analysis was used combining methods, sex and levels of intelligence. A second analysis was employed to throw light on variation in performances from school to school. Differences in non-verbal intelligence scores among the twelve schools (ten children selected at random from each) were adjusted by means of covariance.

Subsidiary Investigation.

One of three chief criteria considered necessary for evaluating i.t.a. was the incidence of backward readers it produced. This was used in a study involving a diagnostic approach to a sample of the backward readers discovered.

RESULTS.

(1) Children (mean age 7·7 years), taught by i.t.a. for three years, were neither superior nor inferior in reading achievements to those taught by t.o. from the outset.

(2) Spelling standards of children who learned to read by i.t.a., were not adversely affected. Girls were significantly better spellers than boys by whichever method they started to read.

(3) i.t.a. produced neither more nor less backward readers than t.o.

(4) Intelligence was a significant factor in determining level of reading achievement. The effectiveness of i.t.a. was neither superior nor inferior in respect to the teaching of bright, dull or average children.

(5) Although there was no significant sex difference in reading attainments, boys seemed to benefit more than girls by i.t.a.

(6) When intelligence was held constant significant variation between schools in reading performances was revealed, indicating the importance of some factor, such as better teaching.

(7) The above results notwithstanding, there was some indication that the consistency of i.t.a. made it easier to read and easier to spell than t.o.

PERSONALITY CHARACTERISTICS AND ATTITUDE TO SCHOOL

BY GEORGE REGAN

(Abstract of thesis submitted for the degree of M.Ed., University of Manchester, 1966)

AIM OF THE RESEARCH.

Attitude to his school was seen as being an index to the satisfactions which the school was affording to the pupil. The investigation set out to discover whether a wide range of personality ' types ' or patterns found equal satisfaction in the same kind of school, or, alternatively, whether different types of schools afforded most satisfaction to different types of personality or, as a further possibility, whether a certain type of personality found equal satisfaction in school, whatever the type of school.

THE SAMPLE AND TESTS USED.

For the main experiment, a sample of 368 second-year secondary school pupils was used. 192 children (98 girls, 94 boys), were drawn from grammar, technical and secondary modern schools and a further 176 children (88 girls, 88 boys) came from a single comprehensive school. The comprehensive school sample was obtained from levels corresponding to the three schools in the ' Tripartite ' group. All the schools were in industrial urban areas in North-West England.

Attitude to school was measured by a twenty-item, Thurstone type scale, constructed by the author. The split-half reliability of the scale was $r=0{\cdot}781$ and a test re-test reliability of $r=0{\cdot}808$ was obtained after an interval of three weeks.

Two personality measures were taken, viz.: (i) The High School Personality Questionnaire (H.S.P.Q.), Forms A and B (R. B. Cattell, 1958). This gives fourteen ' first order ' factors of personality and two ' second order ' factors, anxiety and extraversion. (ii) The Junior Maudsley Personality Inventory (W. D. Furneaux and H. B. Gibson, (1961), which gives two scores an " E " score extraversion), and an " N " score (neuroticism).

ANALYSIS OF THE DATA.

Three techniques were used: (i) Analysis of variance was carried out to determine whether any differences on any of the variables existed between level or type of school or between the sexes. For this analysis only, a random selection of 264 subjects was made from the main sample so as to give twenty-two boys and twenty-two girls in each school or Comprehensive level. (ii) The personality variables were correlated with the attitude to school scale, this being done separately for each school or Comprehensive level. (iii) A factor analysis was carried out using the Principal Components method.

RESULTS.

Analysis of variance showed that both boys and girls in the technical school group had a significantly more favourable mean attitude score than their counterparts in the comprehensive middle group ($P < \cdot 01$). In addition, the technical school boys had a more favourable attitude than the secondary modern and grammar school boys ($P < \cdot 05$). The technical school girls had a signicantly higher mean score than the grammar school girls ($P < \cdot 02$).

The analysis of the scores on the personality factors showed the technical school group to be significantly less anxious than either the secondary modern or grammar school groups, both of the latter being more anxious than their comprehensive counterparts. No anxiety differences were found to be significant within the levels of the comprehensive group, and the range of anxiety scores was narrower in this group.

The technical school group scored significantly higher than either the grammar or secondary modern groups on Factor G (super-ego strength), and lower than the secondary modern group on Factor E (dominance).

The correlation results showed a similar pattern over the different groups, sufficient to justify treating them as one set of results.

Positive correlations with attitude to school appeared for: Factor G (character or super-ego strength), in all six groups (r ranged from 0·261 to 0·593). Factor Q_3 (high self-sentiment), in five out of the six groups, the sixth just missing significance (r ranged from 0·257 to 0·363).

Negative correlations with the attitude scale were found on: Factor E (Dominance), in four out of the six groups, (r ranged from –0·429 to –0·494). Anxiety in all the Tripartite groups but in none of the Comprehensive groups (r ranged from –0·363 to –0·338). The signs in the latter groups were in the same direction, but the correlations failed to achieve significance.

The correlations are moderate or small in size but are significant.

The factor analysis results also showed a similarity in all groups except the secondary modern. In all cases two significant factors were extracted. For the other five groups the factors were: Factor I (accounting for between 28 per cent. and 32 per cent. of the variance), cut across the anxiety *v.* adjustment and extraversion *v.* introversion factors to give an anxiety, introversion *v.* adjustment, extraversion factor. Factor 2 (accounting for between 15 per cent. and 22 per cent. of the variance), distinguished between the adjustment and extraversion aspects of Factor I, and associated favourable attitude to school with adjustment particularly in respect of ego strength, super-ego strength and adventurousness.

In the secondary modern school group the factors were: Factor I (accounting for 28 per cent. of the variance), opposed tough stable, extraversion to anxiety and introversion. Favourable attitude to school was associated with the first pole. Factor 2 (accounting for 20 per cent. of the variance), distinguished between the stability and extraversion factors on Factor 1.

CONCLUSIONS.

None of the data obtained in the experiment could be held to account for the more favourable attitude to the technical school group relative to their contemporaries in the middle comprehensive group. This could either be due to factors outside the scope of the investigation or be a random result arising from the sampling.

Both the correlation and the factor analysis results tend to suggest that there is an ' ideal ' type of personality which makes for good attitude to school and that this type is characterised by self-sufficiency, perseverance, self-control and fastidious individualism. The fact that in the secondary modern school extraversion makes also for favourable attitude could be due to the adoption of a different approach which allows extraverts more opportunities for satisfactions. Alternatively, it may again be a random result arising from sampling.

The lack of differences and general similarity of the results suggested that all the schools are affording greatest satisfactions to a fairly narrow range of personality types and not, as is often claimed, to a wide range of personality patterns.

SOME EFFECTS OF INCREASING THE TEACHERS' KNOWLEDGE OF THEIR PUPILS' SELF PICTURES

By JEAN A. CHADWICK

(*Abstract of thesis submitted for the degree of Master in Education of the University of Manchester,* 1966)

AIM OF THE INVESTIGATION.

This research was designed to investigate whether teachers might, with guidance concerning the characteristics of and the aims for healthy self-conceptual development, be able to incorporate into their role, behaviour believed appropriate to the healthy development of their pupils' self-concepts, and, further, whether such behaviour, in conjunction with knowledge of their pupils' existing self-concepts, might result in changes that were attributable to such behaviour and that had occurred without detriment to academic progress.

The investigation, conducted with secondary school children, bears relationship to an experiment conducted by Staines (1954), with junior school children, in which, although few significant differences in changes occurred between the experimental and control groups, certainty and differentiation about the self were found to increase significantly in the experimental group and areas of the self were identified as points of anchorage.

SUBJECTS.

Three teachers and the three classes of children comprising the 12–13-year age group of a streamed secondary modern girls' school participated in the experiment. The school was in an industrial area and the children came mainly from working-class homes.

INSTRUMENTS.

(*a*) Guidance for the teachers took the form of written notes and any consequently desired discussion.

(*b*) Concepts of pupils' Real, Ideal and Other Selves were measured by a nine-point self-rating test devised by Staines (1954) with a provision for indicating ignorance of any aspect of the self included in the test. Cattell's H.S.P.Q. was also administered to measure general overall anxiety.

(*c*) The N.F.E.R. Intermediate (semi-diagnostic) English Test 1 and the N.F.E.R. Mathematics Test 1 were administered to estimate academic progress.

(*d*) The N.F.E.R. Verbal and Non-Verbal Intelligence Test 1 was used for initial descriptive purposes.

METHOD.

The duration of the experiment covered two school terms: January–March and April–June. Each teacher was the form-mistress for one of the groups and responsible for teaching English and History. Mathematics was taught to all three groups by the same teacher, who was outside the experiment.

The C stream and their form-mistress were experimental throughout the period, the A stream and their form-mistress acted as controls throughout the period, and the B stream and their form-mistress acted as controls for the first half of the period and then became experimental for the second.

All the tests were administered at the beginning, middle and end of the period, except for the intelligence test, which was given at the beginning only. The three teachers also rated their own pupils on the items comprising the self-rating test on all three occasions of testing.

Teachers received information as follows:

After the first testing:

Teacher C—Guidance on the characteristics of and aims for healthy self-conceptual development. Her own pupils' self-ratings. Her own pupils' scores from the Cattell H.S.P.Q. test for anxiety, neurotic trend, extroversion/introversion. Her own pupils' scores on the attainment tests. Her own pupils' intelligence test scores.

Teachers B and A—No information.

After the second testing:

Teacher C—As for the first testing, except for guidance notes and intelligence test scores which she already had.

Teacher B—As for teacher C after the first testing.

Teacher A—No information except for attainment scores which took the place of school examinations.

ASSESSMENT OF THE TEACHERS' ROLES.

This was done through the sampling of lessons during which as much as possible of teacher-comment, except for exposition of lesson content, was recorded. This material was itemised and the items assessed for self-referent content (i.e., referent to pupils' selves) and such content analysed for contribution to self-evaluation and self-knowledge. The assessment and analysis were made by three judges and items where agreement could not be reached were discarded. A reliability test of the judges' assessments, conducted five months later on the contents of two lessons selected at random, yielded reliability figures as follows:

For selection of items considered self-referent:

Judge A	97·07%
Judge B	97·07%
Judge C	97·07%

For detailed analysis of self-referent comments:

Judge A	88·24%
Judge B	84·16%
Judge C	85·97%

RESULTS OF THE INVESTIGATION.

(i) *Teachers' Roles*—χ^2 was used for statistical analysis of teacher-comment and revealed significant differences at the ·01 level between the experimental and control teachers with regard to use of positive and negative self-referent comment and to use of self-referent comment in general. The differences suggested that the experimental teachers had incorporated into their roles the appropriate behaviour.

In conjunction with the other results, however, this finding raised conjecture as to how far the overt aspect of behaviour is effective or representative of the totality of the behaviour of which it is an aspect.

(ii) *Self Ratings.*—Changes in self-ratings were estimated by calculation of significance of difference between means for the groups. Whether based on individual items or on groups of items, significant differences were few and not considered to be related to the experiment.

(iii) *Certainty.*—Changes in certainty about the self were measured by the increase or decrease in use made of the provision in the self-rating test for indicating ignorance of any aspect.

Comparisons between groups and between tests, tested by χ^2, revealed a tendency for all groups to increase in uncertainty about the self as a whole, but the control group, the A stream, although significantly more uncertain than the other two groups at the beginning (·01) (·01) and middle of the period (·01) (·01) was not so at the end ; moreover, while the other two groups' increases were significant in both periods (first half of the period B ·01 ; C ·05 : second half of the period B ·01 ; C ·01) A's increase, although significant in the first (·01) was not in the second.

This pattern of increase was reflected in the Real Self and findings suggested that influence had possibly precipitated into the Ideal Self but not into the Other Self where, it is tentatively suggested, experimental effects may have been operative.

These findings raised conjectures about the influence of various secondary school factors . influence of the 11+ examination and of streaming.

(iv) *Differentiation.*—The facility to see oneself in relative rather than in absolute terms seemed, as measured by use of extreme ratings, to be related to cognitive ability.

(v) *Anchorage.*—Items were ranked for each group at each test for Real and Ideal Self ratings, according to the frequency with which they were rated No. 9 and No. 1 and thus an order of importance or anchorage estimated. Consistency, rather than change or difference, was revealed and areas appearing of greatest importance were those concerning acceptance (especially by the family), values, and physical appearance.

(vi) *Relationship between Anxiety and Real Ideal Self Discrepancy.*—Comparison of High Anxiety and Lower Anxiety scoring children (so designated by their scores on the Cattell H S.P.Q.) on the basis of the proportion of Real/Ideal Self Rating discrepancies that were two or more S.Ds. beyond the mean revealed no significant differences.

(vii) *Differences between Levels of the Self.*—Actual discrepancies between ratings for the different levels of the self were made proportions of possible discrepancies and tests for significance of difference made between groups at each test and between tests for each group. No significant differences were revealed.

(viii) *Academic Progress.*—Rate of academic progress, as assessed by English and Mathematical Attainment tests, did not seem to have suffered any detrimental effect.

REFERENCE.

STAINES, J. W. (1954). *A Psychological and Sociological Investigation of the Self as a Significant Factor in Education.* Unpublished Ph.D. Thesis, University of London.

THE RELIGIOUS ATTITUDE OF SECONDARY MODERN COUNTY SCHOOL PUPILS

By W. P. L. JOHNSON

(*Abstract of a thesis presented for the degree of M.Ed., University of Manchester,* 1966)

PURPOSE OF THE EXPERIMENT.

(*a*) To devise a test to measure the strength and direction of pupils' attitude (i) to religion in general and (ii) to seven main aspects of religion as denoted by the sub-scales mentioned below. (*b*) To give this test to secondary modern school pupils in the first and fourth years of their course. (*c*) To ascertain whether the attitude differed as a result of (i) attendance at Church or Sunday school, (ii) age, (iii) sex, (iv) ability, and (*d*) To view possible results as an indication to improvement in the teaching given at school.

TESTING TECHNIQUES.

A religious attitude scale of forty-two items was constructed, consisting of seven sub-scales, of six items each, concerning attitude to (*a*) God, (*b*) others, (*c*) the Bible, (*d*) religious faith, (*e*) the church, (*f*) school and (*g*) Sunday school.

Insufficient schools were available to allow random sampling, but the group used is believed to be typical of urban secondary modern schools in the Manchester and Stockport area. Starting from 1,440 pupils a stratified sample was selected in respect of the following factors, each at two levels : *Attendance* (at Church or Sunday school) : once per month or more often ; less than once per month. *Age* : eleven plus and fourteen plus. *Sex* : equal numbers of boys and girls. *Ability* : top stream and bottom stream out of four streams. There were thus sixteen categories. A final selection of 640 pupils was made, consisting of forty pupils in each category. These pupils completed the final attitude test. Apart from reference numbers, the responses were anonymous.

The final attitude test was in two parts : A—A self-assessment of attitude to religion, on a five-point scale. B—The attitude scale and its constituent sub-scales.

MAIN RESULTS.

(1) Significant differences in attitude were obtained for the following effects : sub-scales, attendance, age and sex, and between the following interactions : sub-scales × attendance, sub-scales × age, sub-scales × sex, attendance × age, attendance × sex, and sub-scales × attendance × sex × streams. In other cases the differences were insignificant.

(2) The pupils' self-assessments of their attitude to religion were consistent with the scores they obtained on the whole attitude scale. The differences between the means were substantial. This may, therefore, be regarded as a partial check on the validity of the attitude test as a whole.

IMPLICATIONS FOR THE SCHOOL.

(1) We assumed at the outset that an attitude towards religion may be broken down into attitude towards : (i) God, (ii) others, (iii) the Bible, (iv) religious faith, (v) the church, (vi) school, and (vii) Sunday school. The results of the research emphasize the need to recognize the place of each of these aspects in our teaching.

Older pupils have a significantly less favourable attitude to religion than younger, and boys than girls. This confirms the results of Garrity (1960), but is slightly at variance with those of Hyde (1959).

(2) This may indicate the need to relate the teaching to the intellectual and maturational changes of older pupils. Failing this they may find their lessons childish and irrelevant to modern life. This is in line with the findings of Goldman (1962) and Loukes (1961).

REFERENCES

GARRITY, F. D. (1960). *A Study of the Attitude of Some Secondary Modern School Pupils Towards Religious Education*. Unpublished M.Ed. thesis, Manchester.

GOLDMAN, R. J. (1962). *Some Aspects of the Development of Religious Thinking in Childhood and Adolescence*. Unpublished Ph.D. thesis, Birmingham.

HYDE, K. E. (1959). *The Communication of Religious Ideas and Attitudes Among Secondary Modern Children*. Unpublished Ph.D. tnesis, Birmingham.

LOUKES, H. (1961). *Teenage Religion*. London: S.C.M.

THURSTONE, L. L., and CHAVE, E. G. (1929). *The Measurement of Attitudes*. Chicago: University of Chicago Press.

A COMPARISON OF VERTICAL GROUPING AND HORIZONTAL GROUPING IN THE INFANT SCHOOL

BY MARY A. MYCOCK

(*Abstract of thesis submitted for the degree of M.Ed., University of Manchester,* 1966)

STATEMENT OF THE PROBLEM.

In recent years there has been a growing interest in the relationship between school organisation and the development of children, and some large urban infant schools have voluntarily adopted a system of vertical grouping, in which all classes are parallel, each class contains an approximately equal proportion of children of all ages from 5 to 7 years, and the children remain throughout the infant stage in the care of one teacher.

The aim of this investigation was to test the hypothesis that the effects of vertical grouping are more beneficial to children at the infant stage than those of horizontal grouping.

SAMPLE.

The investigation was designed to study selected and matched groups of children in schools which were similar in essential characteristics but which were organised in two different ways, viz., using either a vertical or a horizontal classification of pupils.

Four infant schools A, B, C and D, with school populations of 150–180 children each, were selected. All were characterised by their progressiveness of outlook, informality of regime, purposefulness of work and stability of staffing. Schools A and B (the experimental group) used vertical grouping. Schools C and D (the control group) used horizontal grouping. The schools were paired for locality, size of school, quality of school building and amenities, size of classes, staff ratio, similarity of teaching methods and quality of staffing.

Details of the sampling of children from the twelve classes involved are given below.

METHOD.

A list of the advantages claimed for vertical grouping was compiled and some of the claims used as a basis for the experimental design. These claims were that: (1) there is reduction in admission stress, and increased speed of socialisation for new entrants to school because there are fewer new children in each class and more older people to help them; (2) there is better social adjustment, and a wider range of social interaction, for all children; (3) the lengthened period of association with one teacher makes for better teacher–child relationships and, therefore, greater emotional security; (4) there is elimination of the emotional stress and disturbance involved in 'moving up' since there is no transition from one class to another; (5) there is less anxiety about school work since the organisation ensures continuity of method, and reduces competition and strain by focussing attention on the needs of individual children ; (6) levels of aspiration are raised because of the stimulus of the greater range of opportunities and materials in each classroom ; (7) vocabulary growth is accelerated because of the richer background of experience provided for a mixed age-group; (8) school achievements are better because the classification forces attention upon individual needs, and there is especially more progress in those skills which depend upon language and experience, viz., reading and mathematical skill and understanding.

MEASURING TECHNIQUES.

(1) *Admission stress and speed of socialisation of new entrants.* A time-sampling technique was used to observe and record the behaviour of forty new entrants, aged 5 years, at four intervals during the first month in school. Frequencies were obtained of signs of stress, negative adjustment, positive adjustment.

(2) (*a*) *Social adjustment and* (*b*) *range of social interaction.* A time-sampling technique was used to obtain records of the playground behaviour of a random sample of 144 children (drawn equally from three age groups) in four periods of observation. Frequencies of different types of social behaviour were obtained.

(3) *The effects of lengthened teacher-child relationships.* A sentence-completion test and a drawing test, designed to give insight into teacher-child relationships by providing comparative data on the child's attitudes to his mother and his teacher, were administered to matched groups in a sample of 154 seven-year-olds.

(4) *The effects of eliminating transition from class to class* Class teachers made two registrations of the Bristol Social Adjustment Guides (Stott 1958) for each child in a total sample of 112 children aged 6 years. Report 1 was made two weeks before the end of the Summer term. Report 2 was made two weeks after the beginning of the Autumn term. The completed Guides were scored by the investigator and total adjustment scores obtained.

(5) *Anxiety about school work.* The Test Anxiety Scale for Children (Sarason, 1960) was administered individually to 165 children aged 7 years, in their third year in school.

(6) *Levels of aspiration.* Seventy-two children (equally drawn from three age groups), paired for age, sex, and intelligence, were given a peg-fitting test, in which a number of timed trials yielded a series of attainment scores and goal discrepancy scores.

(7) *Size of vocabulary.* 108 seven-year-old children, paired for age, sex and intelligence, were tested individually using Watt's Vocabulary Test for Young Children (1945).

(8) *School achievements.* 108 seven-year-old children, paired for age, sex and intelligence, were given tests for (i) Reading accuracy, comprehension and rate, using the Neale Analysis of Reading Ability, and (ii) Mathematical skill and understanding, using two Piaget-type tests.

All tests were carried out and data collected within a period of one school year and two weeks.

RESULTS.

No significant differences between the two groups were found in (2) (*a*) Social adjustment ; (5) Anxiety about school work ; (7) Size of vocabulary ; and (8) School achievements in reading and mathematical skill and understanding.

Significant differences in favour of the experimental group were found in : (1) Admission stress and speed of socialisation of new entrants ; and (4) Effects of transition from class to class ; but evidence suggested that such differences might be of limited duration.

Significant differences in favour of the experimental group were found in : (2) (*b*) Range of social interaction ; (2) Effects of lengthened teacher-child relationships, and (6) Levels of aspiration.

INTERPRETATION OF RESULTS.

The original hypothesis was not supported in the results connected with social adjustment, anxiety about school work, size of vocabulary, and level of school achievements. There was some evidence that in reading, slow learning children benefited more from a policy of vertical grouping whilst bright children profited more from a policy of horizontal grouping.

It was found that vertical grouping facilitated adjustment more effectively in the early stages but evidence suggested that such gains might be of limited duration, although nervous immature children appeared to find adjustment more difficult in a horizontal organisation than in a vertical organisation. In transition from class to class, the evidence suggested that regression and emotional disturbance caused by class transfer might be of a temporary nature. Judged by the control group in this sample, the check would appear to be a temporary one, and once the subjects have become familiar with the new situation, it seems probable that they regain lost ground and display social behaviour very similar to that found in the experimental group.

The results indicated that children in a vertical organistaion had a wider range of immediate and available roles, developed better work attitudes, and had a closer, warmer relationship with the teacher which implied greater emotional security and therefore, stability. The findings emphasised the crucial importance of the quality of the teacher in a vertical organisation.

THE RELATIONSHIP BETWEEN MOTIVATION, PERSONALITY, INTELLIGENCE AND SCHOOL ATTAINMENT IN A SECONDARY MODERN SCHOOL

By MARJORIE E. AINSWORTH

(Abstract of a thesis submitted for the M.Ed. degree of the University of Manchester, 1966)

AIM.

The aim of the investigation was to assess the value of two comprehensive tests for children, of personality and motivation, the criteria of usefulness being a consideration of the results obtained from them in the following experimental fields :

(1) To assess cross-cultural differences between British and American children in personality and motivation.

(2) To investigate the relationship between intelligence, attainment and personality factors.

(3) To investigate the relationship between intelligence, attainment and motivation factors.

(4) To investigate the relationship between sex, personality and the hierarchy of streams within a school.

(5) To investigate the relationship between sex, motivation factors and the hierarchy of streams within a school.

(6) To assess the suitability of the High School Personality Questionnaire for the individual assessment of children.

METHOD.

230 children from the eight streams in the first year of a mixed secondary modern school completed the I.P.A.T. Culture Free Intelligence Test, Forms A and B, the I.P.A.T. High School Personality Questionnaire, Forms A and B, and the four sub-tests of the I.P.A.T. School Motivation Analysis Test. Criteria of

attainment were the scores obtained in the annual internal school examinations in English, Mathematics, Science, History, Geography and Art. Comparisons of means and standard deviations on fourteen personality factors, were made with similar reported American and British results in connection with Experiment 1. Product-moment correlation coefficients were calculated between Intelligence Quotient and each of the six measures of school attainment, between Intelligence Quotient and each of the fourteen personality factors, between each of the fourteen personality factors and each of the measures of school attainment, between Intelligence Quotient and each of the fifteen motivation factors, and between each of the fifteen motivation factors and each of the measures of school attainment, in connection with Experiments 2 and 3. Three-way analyses of variance between sex, school stream and personality factors, and sex, school stream and motivation factors were calculated in connection with Experiments 4 and 5. Two-way analyses of variance were calculated between the questions representing each of the fourteen personality factors measured by the High School Personality Questionnaire and the responses of a random sample of twenty boys, as a check on the reliability of the factors and the ability of the items contributing to them to discriminate between individuals, in connection with Experiment 6.

RESULTS.

Experiment 1. Amercian children were shown by the High School Personality Questionnaire to be very significantly more friendly, more conscientious and less dominant than British children.

Experiment 2. Intelligence Quotient correlated positively and significantly at the 1 per cent. level with all the measures of school attainment except Art. None of the fourteen personality factors correlated significantly differently from zero with Intelligence Quotient, with the exception of Factor B, which is also a measure of intelligence. Conscientiousness and self-sufficiency were the most promising factors as predictors of school attainment, although none of the correlations between attainment and personality factors, approached the magnitude of thses obtained between Intelligence Quotient and school attainment.

Experiment 3. Assertion was the only one of the fifteen motivation factors correlating significantly with Intelligence Quotient, and that negatively at the 5 per cent. level. The super-ego sentiment correlated positively and significantly with all the measures of school attainment.

Experiment 4. The analysis of variance revealed a second-order interaction effect, significant at the 5 per cent. level between sex, stream and personality factors. Separate analyses of variance were, therefore, calculated for boys and girls, to isolate the interaction of streams and personality factors for each sex. Where significant differences in personality factors between school streams were found, in no case, either for boys or girls, did the differences follow the assumed hierarchical order of the streams. There is some evidence that the brighter girls are more friendly than the duller girls, and that the duller girls are more dominant than the brighter girls.

Experiment 5. Girls were shown to be significantly more motivated by sensuality, narcissism, sex and religious sentiment than boys, whereas boys were shown to be more constructive, more patriotic and more likely to indulge in play-phantasy than girls. The super-ego sentiment most nearly differentiated between the school streams, the two top groups scoring significantly higher on this factor than the remaining streams. The lowest groups were shown to be most assertive, aggressive and patriotic. The top group was shown to be the most narcissistic, the most acquisitive, the most religious, to indulge most in play-phantasy and to have the most highly developed super-ego sentiment.

Experiment 6. The ability of the test items representing personality factors to discriminate between individuals is doubtful in the case of most of the factors. In the analysis of variance check, only five of the fourteen factors discriminated at a significant level between individuals on both or one of the two forms of the test. The reliabilities calculated from the analyses of variance are much lower than would normally be expected from a questionnaire. Those found for four of the factors are of the same order as those quoted in the test handbook, and would presumably be regarded as adequate by the authors of the test, who insist that reliabilities on a test based on factor validity will necessarily be lower than those found for tests based on the ideal of homogeneity. The very low and negative reliabilities found here on some of the factors may have been improved by a more standardised test situation in which all the subjects were tested in one group under a single administrator, as the authors of the test comment that the High School Personality Questionnaire is sensitive to test procedure and the motivation of the subjects in a group situation.

CONCLUSION.

These two comprehensive tests for children, of personality and motivation, have proved to be easily administered and objectively scored tests, covering a wide area of personality and motivation factors. Several of the tested factors aid in the prediction of school attainment, and discriminate between the sexes, between status groups and between culture groups. The high face validity of the reported significant differences, and close correspondence with other reported results,lend weight to the view that these tests are suitable instruments for use in group studies and surveys. It is felt that the High School Personality Questionnaire would be of doubtful value for the individual assessment of children as the test items, with a few exceptions, do not adequately discrimintate between indviivduals.

FACTORS AFFECTING THE COMPOSITION PERFORMANCE OF TEN-YEAR-OLD CHILDREN

BY DEREK SHARPLES

(*Abstract of thesis presented for the M.Ed. degree of the University of Manchester*, 1966)

AIM.

Previous research has indicated factors related to attainment in primary schools, but has not explored these in relation to reliable estimates of performance in composition. The effects of schooling and maturation have been noted, but the interaction of these effects with other factors has not been examined. The present study considers seventeen factors related to attainment and development in composition writing amongst children aged 10+.

METHOD.

A battery of compositions and tests was completed by a sample of 127 children drawn from five primary schools of varied social background, as established by means of the juror's index.

The factors examined were as follows:

Composition performance. Nine compositions were written in response to narrative, descriptive and expository titles. Marking was by impression using four markers, as described by Wiseman (1949); marks were standardised for each title to avoid weighting and pooled to yield a mean score for each child.

Composition development. Three composition titles were repeated after a period of twelve months. The responses from both occasions were marked together in single title batches by impression methods. This technique facilitated a comparison of the performances on each occasion and was used to assess schooling and maturation.

Attainment. Standardised tests of verbal intelligence and English were employed. A comprehension test of Watts (1944) was administered after reliability and validity were established.

Interests. Child interests were established by an instrument which invited children to indicate their response to titles such as " My favourite game,"" An interesting hobby," etc.

Attitude. A Thurstone type instrument of twelve items was constructed to indicate attitudes of children towards composition writing.

Personality. A revised form of the Junior Eysenck Personality Inventory was administered.

Socio-culture. A test suggested by Campbell (1952) was constructed to indicate the degree to which children were familiar with favourable cultural and social experiences.

Appreciation. Children were required to rank five compositions; their rankings were correlated with those of a group of teachers to indicate the degree to which the child's appraisal concurred with that of teachers.

Schooling. Data were obtained to describe the physical environment of, and teacher attitude and methods of teaching composition in, the sample schools.

Environment. Data were obtained of social class, family size, age and sex of each child in the sample.

All tests were administered by class teachers in the schools following detailed procedural instructions or employing recorded directions standard to all schools.

RESULTS.

Analysis of variance showed significant differences between the scores of social classes, groups of differing family size and sexes; ' t ' tests among the means showed the differences to be in favour of higher social classes, small families and girls.

Three composition performance groups were defined by reference to composition scores, viz., successful, average, and unsuccessful performance. Analysis of variance of the test scores of these groups showed significant differences between groups in verbal intelligence, English, comprehension and socio-culture, in favour of the successful groups. There were no significant differences in personality, attitude or appreciation scores.

Analysis of variance of the mean composition scores of the schools showed consistently significant differences in favour of three of the sample schools as compared with the other two. The three superior schools were those in areas of high ' j ' index, having a larger proportion of children from higher social classes than the inferior schools. Differences in test of verbal intelligence, English and comprehension were in favour of the schools with the highest composition scores. There were significant differences in attitude scores in favour of the school displaying the lowest mean composition score. No significant differences between schools were found in other areas, although teacher method appeared to be related to the abilities of the pupils.

Anaylsis of co-variance of the repeated titles showed that there were significant differences in the development of composition performance between sexes, social classes and attainment levels, but not between schools. ' t ' tests showed that girls, higher social classes and more able children make significantly greater improvement in composition performance. In all schools attitudes fell, but significantly smaller falls were found in schools which appeared to be least subject to parental pressure and more permissive in approach to composition teaching.

Correlations between significant factors and composition are shown in Table 1.

Table 2 shows correlations between composition scores. The correlations suggest that although narrative titles do not offer the highest marks to all children (Bilborough, 1957), yet they give the most reliable prediction of overall composition performance. Other predictors appear to be less economic and are no more promising.

TABLE 1

CORRELATIONS BETWEEN SIGNIFICANT FACTORS AND COMPOSITION.

Test	(*b*)	(*c*)	(*d*)	(*e*)	(*f*)
(*a*) Composition	·719*	·756*	·310*	·302*	·251*
(*b*) Verbal Intelligence	—	·861*	·801*	·068	·041
(*c*) English		—	·682*	·134	·135
(*d*) Comprehension			—	·293*	·293*
(*e*) Socio-culture				—	·089
(*f*) Appreciation					—

* Significant at 1 per cent. Others not significant.

TABLE 2

CORRELATIONS BETWEEN SCORES ON DIFFERENT KINDS OF COMPOSITION.

Component	Pooled Score	Pooled Score Less Component
Narrative	·963	·741
Description	·922	·586
Exposition	·812	·570

All correlations significant at 1 per cent.

CONCLUSIONS.

The conclusions from the study indicate some of the factors affecting composition performance amongst children of average and above average intelligence, and are limited by the bias of the sample towards higher social classes.

(1) Girls, children from higher social classes and from small families are most successful in composition writing. Whilst composition performance improves with age, there are no significant differences between age groups within classes.

(2) Children superior in composition writing are also superior in verbal intelligence, English attainment and comprehension. There is no evidence to suggest that personality or interests are related to superiority.

(3) Children superior in composition writing hold less favourable attitudes towards composition than less successful children, probably as a result of schooling experience.

(4) Favourable cultural experiences are related to success in composition writing among children who are able to read.

(5) Children show significant improvement in composition performance during the final year in the primary school. The schooling experiences favour children already successful, and differ in their effects on attitudes towards composition. When initial differences are allowed for there are no significant differences between the gains of different schools.

(6) The ability to evaluate written language is related to the ability to write composition.

(7) A limited battery of narrative titles is a powerful predictor of overall performance level in composition.

REFERENCES.

BILBOROUGH, H. (1957). *The Effect of the Title on Performance in Composition.* Unpublished M.Ed. thesis, Manchester.

CAMPBELL, W. J. (1952). The influence of home environment on the educational progress of selective school children. *Brit. J. Educ. Psychol.*, 22, 89-100.

WATTS, A. F. (1944). *The Language and Mental Development of Children.* London: Harrap, p. 354.

WISEMAN, S. (1949). The marking of English compositions in grammar school selection. *Brit. J. Educ. Psychol.*, 19, 200-209.

PERSONALITY FACTORS AND ATTITUDES OF MATURE TRAINING COLLEGE STUDENTS

BY EDITH SOLOMAN

(*Abstract of thesis submitted for the degree of M.Ed. in the University of Manchester,* 1965)

AIMS

The research had two main aims—to investigate the personality factors and attitudes of mature student-teachers in relation to success in a training course, and to establish the validity against established questionnaires of a new series of objective personality tests.

SAMPLE

The sample consisted of 155 students—seventy-eight men and seventy-seven women—aged from 20 to 53 years, attending a day training college in the West Riding of Yorkshire.

METHOD

(1) 105 variables of investigation and five variables of criteria were obtained as follows:

(*a*) *Biographical Data,* viz., age, marital status, parenthood, sex and interests. (Arts or Science interests were classified according to the subjects chosen for the main academic study during the course): 5 variables.

(*b*) *Questionnaires,* viz., Cattell's 16 Personality Factor Questionnaire, Eysenck's Maudsley Personality Inventory, and Oliver's Survey of Opinions about Education: 25 variables.

(*c*) *The Objective Test Battery,* based on research by Cattell at the University of Illinois, comprising 26 tests: 73 variables.

(*d*) *Ratings,* viz., Ratings of Extraversion and Anxiety made on each subject by tutors at the college. (Descriptions of the two dimensions on which to base their assessments were formulated on primary factor data, in the 16 P.F. Handbook): 2 variables.

(*e*) *Criteria,* viz., grades earned by students in the Final Teaching Practice; their final assessments in the Theory of Education, the Academic Main Study, the Academic Subsidiary Study, and a summed score of all grades labelled "Final Certificate Award": 5 variables.

The tests were incorporated into the Education course at the college and administered at convenient times over a period of four weeks. The total programme of tests took four hours to complete.

(2) (*a*) Means and standard deviations, and product-moment correlations of 110 variables were calculated by computer. Significant correlation coefficients were

extracted for analysis and interpretations. (*b*) All 5,995 correlations were submitted to a further computer programme, which carried out a principal components analysis, followed by a varimax (orthogonal) rotation of factors with latent roots greater than unity.

RESULTS

(1) Table 1 reveals the variables which significantly correlated with the criteria (asterisks show results at the 1 per cent. level of significance).

TABLE 1

INTERCORRELATIONS.

	Practical Teaching	Theory of Education	Academic Main Study	Subsidiary Study	Final Certificate Award
BIOGRAPHICAL :					
Age		243*	175		
Marital Status		231*			
Parenthood		197			
Female Sex	230*	263*			184
Arts Interests	164				
QUESTIONNAIRES :					
Intelligence B (16 P.F.)		189			185
Unconventionality M (16 P.F.)		182			
Liberalism Q1 (16 P.F.)		227*	194		191
Tenseness Q4 (16 P.F.)	159				
Radicalism (16 P.F.)			174		
Radicalism in Education	164	157			167
OBJECTIVE TESTS :					
Speed in Addition		211*		179	226*
Long-term Aspirations				177	
Speed in Decision		250*			
Impulsiveness		—205*			
General Fluency (Item 1)		221*			
General Fluency (Item 2)		160			
Fluency about Self (Item 3)		175		158	169
Fluency about Self (Item 4)		158			
Annoyability				—209*	—183
Persistence in Speed of Tapping	187				
Persistence in Speed of Addition (difficult task)				238*	
Lack of Persistence in Speed of Addition (easy task)			170	244*	201
Fluency of Pleasant Ideas	207*				
High Embarrassment			—197		
Low Opinion of Fairness of Test				176	
RATINGS :					
Self-rating of Extraversion					169
Tutor-rating of Extraversion	250*		171		252*
Tutor-rating of Anxiety			—159	—200	

(2) Analysis of further information included the connections between biographical and questionnaire variables; biographical and objective test variables; biographical and rating variables; questionnaire and rating variables; objective tests

and rating variables. Generally speaking, the results appeared to be in expected directions, with one or two interesting exceptions, e g., that neither women nor parents in this sample were more tenderminded in their attitudes to education. All the criteria variables correlated highly with each other.

(3) Correlations between objective tests and questionnaires were, generally, low. Fifty of the seventy-one objective test variables correlated significantly with questionnaire variables. The number of these associations for each objective test variable ranged from 1 to 17 (average 3 ·7).

Three of the variables showing no significant connection with questionnaires correlated with criteria variables of academic achievement.

(4) The factor analysis revealed twelve major factors underlying the personality traits and attitudes of this sample of mature teacher-trainees. They were labelled and interpreted as follows :

(*a*) Factors related to the major second-order personality questionnaire variables were : Emotionality, Extraversion, Tough-mindedness and Radicalism.

(*b*) Factors related to established objective test factors were : Smartness, Apathy, Dourness and Nonconformity.

(*c*) Factors related to the criteria and biographical variables were : Achievement and Maturity.

(*d*) Factors related to additional personality patterns were : Fluency about self, and general fluency.

MAIN CONCLUSIONS

(1) The successful teacher is a woman who appears to have a certain level of tenseness of personality which would possibly result in a ' live-wire ' quality in the classroom, and an alertness to the children's needs. She is characterised by persistence in speed of personal tempo, which may enable her to maintain this alertness in spite of difficulties and boring routine and the fluency of her pleasant ideas would probably be used to make the children's learning enjoyable. The teacher with ability in this sample has radical attitudes to education, which would probably result in progressive activities in school, and she is considered by her tutors to be an extrovert, i.e., showing a marked warmth, sociability and cheerfulness. Her interests are in the Arts subjects, and her attainment in these is high.

(2) Academic achievement and personality traits significantly associated with it, suggests that originality, liberal, progressive and critical ideas and attitudes play as large a part as intelligence at this level of study. Other influences include speed (of decision and addition) ; fluency of ideas ; persistence and criticalness ; and low levels of annoyability and embarrassment. Tutors rate the academic achiever here as being both stable and extraverted ; and success in the academic aspects of the course was affected positively by age. Older students, women, those who were married and parents reported achievement in The Theory of Education especially. Expressed in terms of the second-order dimensions of personality, success on a teacher -training course seems to depend on stability, extraversion and radicalism.

(3) The significant relationships between a number of the objective tests and several of the questionnaire variables, appear to be an encouraging initial attempt in the validation of the tests, and certain tests appear to justify their inclusion in a personality measurement programme.

CRITICAL NOTICES

THE EDUCATION OF THE GIFTED

BRANCH, M., and CASH, A. (1966). *Gifted Children*. London : Souvenir Press, pp. 222, 30s. net.

BY CYRIL BURT.

This book belongs to a type of publication which is appearing in increasing numbers, what may be called popularized psychology—well-intentioned efforts to interest parents, teachers, and the general public in what are supposed to be the latest results of psychological speculation and research. Usually the authors are neither psychologists nor educationists. As a result, such books commonly receive only scant notice in professional journals—more often indeed, no mention whatsoever. All too frequently, however, the attempt at popularization means a tendency towards excessive simplification, vague and sweeping generalization, and a relish for theories which are novel and startling rather than firmly established and logically expressed. It thus tends to bring the whole subject into disrepute among the very people to whom the writers are seeking to appeal.

The problem of the gifted child, as Professor Lauwerys reminds us in his Preface to the present volume, is one on which a vast literature, largely technical and statistical in character, is already available. Yet it is a problem which, in current controversies about educational reorganization and selection, is either grossly misinterpreted or totally ignored. One would expect, therefore, that every educationist would welcome a book which endeavoured to bring the subject to the notice of the general public. Neverthless, the criticisms which this particular work has aroused in the educational press, and the numerous questions and expressions of dissent that I myself have received, plainly indicate that (to put it mildly) the authors have evoked more doubts than they have solved. On the other hand, the columns of *The Times* and of other daily papers have given a good deal of publicity to some of the proposals which Mrs. Branch and Mr. Cash have advocated. Accordingly, in view of the confusion which still seems to surround the whole subject, it may be worth while to examine in some detail the views which they have put forward, and the facts and theories on which their arguments are based ; and then to consider the various practical measures which they wish to recommend.

Mrs. Branch, so the publishers inform us, is a former personnel manager, who has spent much of her time on voluntary social work both in this country and in Germany, and, having taken a diploma in mental health, is now a psychiatric social worker at a London hospital Mr. Cash is a journalist and script-writer for films, and a co-author of the books and lyrics for a new musical : in addition, we are told, he worked on the B.B.C. film, *The Exceptional Child*. The book itself falls in to two overlapping parts.

The second part was apparently the first to be written, and has not, I think, received the attention it deserves. It deals with the special problems presented by the highly intelligent individuals who, we are told, are not infrequently to be found among children diagnosed as suffering from "some kind of damage or immature development of the brain." Three main groups are described, namely, those who are severely handicapped (1) by gross physical defects (such as are associated with cerebral palsy), (2) by limited mental disabilities (such as dyslexia), or (3) by various types of ' emotional immaturity ' (broadly classified into two sub-groups—the ' maladjusted ' and the ' inconsequential '). Since these chapters are written in the first person singular, we may, I imagine, infer that they are the work of Mrs. Branch. Her writing is by ni means easy to read and is often disfigured by slips that have escaped correction. Her neurological and pathological explanations are rather naive and reveal a good deal of confusion on technical points. And nowhere does she tell us what proportion of the ' brain damaged cases ' come into the category of ' high I.Q. or near-genius.' But, whether few or many, they form a tragic group whose very existence has hitherto passed almost unnoticed; and parents in particular should be grateful to the authors for drawing attention to the special needs and capabilities of these afflicted youngsters. The case-histories (unfortunately all too few, only one or two in each of the main groups) show what can be done if the child's special abilities are recognized at an early stage, and how hopeless the outlook usually becomes when left until too late.

A couple of examples may help to bring home the type of case for which Mrs. Branch is appealing. Mary H., aged 18½, was a student at a teacher training college. She had a way of swallowing large doses of barbiturates in public places—a railway station or a church, just enough to set people talking of suicide, but not enough to prevent her being quickly revived with a dash of cold water and a cup of black coffee. She was admitted as an in-patient. " She repeatedly asked for reassurance," (writes Mrs. Branch) " and I reassured her, explaining that, if she would co-operate, she would be able to resume college the following year . . . Unhappily her cry for help had come too late." There were several relapses; and a few weeks later the girl left hospital against medical advice. No more has been heard of her since. " Was there," asks Mrs. Branch pathetically, " anything I could have done? " We are told nothing of the diagnosis or treatment given by the psychiatrist, simply that " she clearly was a gifted girl." Mrs. Branch classifies her as an ' inconsequential '; and explains that " the man who has done most in this field is Dr. D. H. Stott, who postulates that the mothers of inconsequential children may have had troubled pregnancies and that the children are, in fact, slightly brain damaged." It appears, however, that before deciding to become a teacher, the girl had worked as a shop assistant and in similar low-grade jobs; and one wonders whether a partial cause may not have been that the high aspirations which she cherished were constantly frustrated because she was *not* sufficiently gifted for the academic work required.

By way of contrast, there is the case of a girl who undoubtedly was ' brain-damaged '—a spastic child who all her life has been confined to a wheel-chair. The story is told in the mother's own words, and is a most moving and illuminating instance of what is possible with a case that seems beyond all hope. At the age of 3, ' Joanna ' was still unable to speak, and the doctor declared that she was mentally defective. When she was 7, she was sent to a day spastic unit, where " a teacher who was a natural genius " taught her to paint, model, and read. After that, her mother took the girl in hand, sitting beside her for hours, reading books borrowed from libraries, seeking to interest her in concrete subjects—art, architecture, music, answering her questions (the child's speech was so indistinct that they had to use letters to find the key-word of each problem), and finally giving her a wireless set which she could switch on for herself. At the age of 12 she was accepted by the Delarue school for spastics, but not permitted to attend the O-level class. The mother, however, pleaded; and very slowly Joanna learned to type. " Suddenly, it all began to pour out" (says the mother) "first in letters to me in the lively style of our discussions at home, and then in her essays at school . . . Now she has five O-levels; is doing two A-levels, and an extra O in Spanish."

Working with medical specialists, Mrs. Branch very naturally is convinced that the basic problem is physiological. We are given a brief, rather technical, and not very accurate account of the different effects of damage to the basal ganglia, the cerebellum, and the cerebral cortex. We are told how backwardness in reading results from brain injury, and how damage to other parts " may be death to the possibility of achieving any mathematical skill." In a concluding chapter, we hear of a " theory, well supported by a growing body of empirical discovery, that we are all victims of such injury . . . a brain trauma, so infinitesimal that no instrument exists sensitive enough to reveal it." The injury, it is said, is due to " the placenta's deterioration during the last few months of pregnancy," which impairs the flow of oxygen. The result is a " neutronal (*sic*) damage " (i.e., damage to the neurones of the brain) which prevents most of us from " making full use of our gifts." This explains why " brilliance and high achievement would seem, in many cases, to go hand-in-hand with premature birth "—a point which, so Mrs. Branch considers, is now ripe for research by those who are interested in the psychology of genius. " Eugenesists (*sic*) would put in a strong claim for the hereditary factor ";

but, as we learn on another page, " heredity is magnified " by such writers. Mrs. Branch then goes on to describe the " decompression experiments " carried out by Dr. O. S. Heynes (Professor of Gynaecology at the University of Witwatersrand). " Decompression," she explains, " is a method of introducing oxygen into the bloodstream of pregnant women " by means of " a special machine " (described in detail). The babies born were kept under observation for at least two years, and repeatedly tested : " hardly without exception, these babies were advanced in every way," and one of the mothers, it is said, was herself grossly defective with an I.Q. of only 55. Mrs. Branch concludes that, " even if decompression is not the whole answer to the question of releasing the full potential of human brilliance, the method should receive urgent attention."

This account has given rise to numerous inquiries from parents and others. It may, I think, be said that, so far as the evidence at present available indicates, the procedure referred to (and recently introduced at one or two hospitals in this country) may well diminish the risk of hypoxia in cases where it seems likely. In that way it would certainly help, as Mrs. Branch observes, to reduce the number of spastic cases. But the wider claims seem extremely dubious. The incidence of prematurity among geniuses of the past and among children with high I.Qs. has in fact been investigated by Havelock Ellis and by Terman : their figures suggest that, if anything, it is slightly *less* frequent among the highly gifted. Most investigators seem now agreed that, on the whole, children born prematurely are more liable to be mentally subnormal than children born approximately at the normal term, and that the gravity of the condition increases with the degree of prematurity. However, the problem has proved more complex than is ordinarily supposed : first, it is important to discriminate between ' true ' and ' false ' prematurity ; secondly, it is not always easy to distinguish the effects of prematurity *per se* (i.e., premature exposure to the exacting requirements of an extra-uterine environment) and the effects of the prior pathological processes which were presumably responsible for the premature birth.

Mrs. Branch's case-histories and the inferences drawn are based on her hospital work ; and it is this record of first-hand experience, described from the viewpoint of a social worker, that forms the most original and instructive section of the book. Yet, oddly enough, neither Professor Lauwerys's Preface, nor the publisher's summary of the contents as set forh inside the jacket, makes any mention of these interesting chapters The publishers tell us that " in this book the authors have made a careful study of a unique problem—how can we recognize exceptional intelligence, nad ensure that it is not wasted ; " and Professor Lauwerys, whose century mankind needs, badly and urgently, the contribution of the gifted . .

century makind needs, badly and rugently, the contribution of the gifted . . .

Teachers need to know much more about giftedness, high ability, unusual talent—in short, about a kind of genius ; " and he commends the book as " breaking new ground."

These, in fact, are the topics discussed in the first half of the book. By ' giftedness,' so the authors inform us, they mean " something more than mere brilliance : the problem is—what to do with the youngster whose intelligence is such that he can be assessed as a potential genius." This section is couched in a different and far more colloquial style, and presumably comes from the pen of Mr. Cash. The headings are in contemporary journalese : " All the Colours of the Rainbow," " Little Lamb who Made Thee ? " " First Person Singular," and so on, until we reach " The Whole Child : A Postscript." After this there is an abrupt transition to the chapters on dyslexia, cerebral palsy and the physically handicapped, described above. Possibly the publishers felt that a book for the general reader which dealt solely or chiefly with what would ordinarily be regarded as medical cases would have little sale by itself, and therefore, pressed for a more comprehensive survey and a more popular treatment. In any case these opening chapters seem to have been prefixed as an after-thought, very hastily prepared and still more hastily written.

The authors begin by saying they have " followed every lead they were given," and add that, if they have missed any important work in this field, they hope readers will inform them, " so that corrections can be made in future editions." The most striking omission is the absence of any reference to scientific researches carried out during the present century in this country, and published in this and other psychological journals and in various official reports. Thus, in discussing " the question of selection " no mention is made of the investigations undertaken by the National Foundation for Educational Research. Moreover, it is a little bewildering to find that the theoretical views put forward in the first part are often contradicted by those expressed in the second.

In gathering material for their opening chapters, the authors appear to have relied mainly on two unacknowledged sources. Mrs. Branch, I understand, approached Mr. Victor Serebriakoff, the secretary of a society which at that time was considering the possibility of starting one or more experimental schools similar to those established in the U.S.A. by the Association for Gifted Children. Fresh from a visit overseas, he generously put at her disposal his collection of relevant books and other literature, chiefly American, which seems to have formed the basis of the ' Bibliography.' Secondly, free use has evidently been made of the recent issue of *The Yearbook of Education*, entitled *The Gifted Child*. The introductory chapter, as well as a good deal of the later ones, follows almost verbatim much of the ' General Introduction ' to the *Yearbook*. Both chapters have the same heading— ' Who are the Gifted ? ' (except for an ungrammatical change of ' Who ' to ' What ') ; both start off with the same quotation, " The first duty which the gods have placed upon rulers is to select all the children of gold," etc. ; but the present volume adds, " Thus wrote Plato in *The Republic*." (In point of fact Plato did not write the words transcribed ; they were, as a footnote explained, merely a paraphrase). Over the page we come to an epitome of Galton's methods and views, again borrowed almost word for word (including the quoted extracts) from the following pages in the *Yearbook*. Binet's views are then discussed, with comments manifestly copied from the *Yearbook*, though one or two of the remarks are applied to Binet by mistake. The *Yearbook* selects Norbert Wiener as a typical instance of a ' highly gifted child ' who was sadly out of place even in the upper forms of his high school ; and cites a passage from his autobiography describing how " a deep estrangement fell upon him " etc. ; the same example, with the same quotation, and even the same omission in the middle of the quotation, appears in *Gifted Children*.

Nowhere in these chapters is any acknowledgement made as to the sources of all this information. Naturally, a book for the general reader should not be loaded with references or footnotes. Nevertheless, when (as so often happens) a popular writer reproduces the conclusions of some particular scientist without the usual acknowledgements, a good deal of misunderstanding is apt to result. Sometimes the scientist is subsequently accused of borrowing his ideas from the more popular writer, who has become better known to a wider public ; sometimes the reviewers and critics attribute the mistakes of the popularizer to the originator of the views in question.

These earlier chapters lay far more stress on the importance of a " modern knowledge of genetics." But the writers' own exposition of " modern genetics " and its bearing on educational problems, is full of naive errors and misconceptions. No wonder egalitarian critics, in discussing the authors' arguments, claim that it shows " how completely the believers in innate mental differences misapply and falsify the principles deduced by neo-Mednelian geneticists for physical differences."

A still more serious confusion arises from the authors' definition of the type of child whose special needs they are proposing to discuss, namely, what they call the ' potential genius.' In the *Yearbook* the first table, based on Galton's figures, classifies children into various categories, and gives borderlines in terms of (*a*) the I.Q., and (*b*) the corresponding percentage. Galton's definition of genius, it is explained, implied a borderline of 140 I.Q. : this, according to the table, cuts off the top 2 per thousand, or thereabouts. On the other hand, the earliest British studies were concerned with a much larger group, namely, the top 2 per cent., that being the approximate proportion in each age-group who, in those days, obtained a junior county scholarship to a ' secondary ' (grammar) school. Thus, a fairly sharp distinction was drawn between two types and definitions of ' gifted children '—the ' moderately supernormal ' and the ' highly supernormal.' In this country most of the research has been concentrated on the former group, and a good deal is now known about their subsequent careers and the reasons for their success or failure. Nevertheless, it was said, the special problems raised by the higher type are far more perplexing and have hitherto passed almost unrecognized. As a group they are as far above the average child in the grammar school as he is above the average of the general population. These are pupils who, like Norbert Wiener and others, are most likely to suffer from frustration and intellectual maladjustment; and yet it is to these that we must look for our future leaders not only in literature, science, and technology, but also in industry and politics.

Now from the definitions given us by Mrs. Branch and Mr. Cash, it is not at all clear which of these two types they have in mind. They begin by telling us that by gifted children they mean "the vital 2 per cent. of the population," of whose needs, they say, "we are so far comparatively ignorant." But they end the chapter by telling us that their "borderline is an I.Q. of 140+." And both these incompatible definitions are repeated in different parts of the book.

We then have a chapter of interviews with bright boys and girls (I.Qs. not stated) which seeks to show how, in such cases, "boredom, under-achievement, and a self-destructive journey to delinquency is a not uncommon progression," ending with the story of an ex-Borstal boy (illegitimate) who eventually obtained a second-class degree at a provincial university—hardly an example of a 'potential genius.' The next chapter contains half-a-dozen specimen compositions, chiefly poems. Here, we are told, the fantasy-life of the gifted youngsters will provide the "enlightened parent with clear indications of the child's unsuspected feelings of frustration and secret misery." We are offered interpretative comments based on "the use of symbolism in childish fantasy." Susan, aged 8, writes a poem about a November 'guy' flung on to the bonfire, while all around

> Lots of children are having fun
> With fireworks that bang just like a gun;

in the morning there is nothing left but "dust and ashes." "This poem," it is said, "written no doubt straight out of a mood, is a startling revelation of the awareness of pain and unrequitement, while everyone else is having fun." And of another composition (about a tiny elf who escaped from a jar of mistletoe-berries into which a girl had dropped him, and safely returned to fairy land) the authors ask: "Isn't this to do with love? . . . What could be plainer than this description of feeling unimportant, unloved, and overlooked? Don't we kiss under mistletoe?" The interpretation is plain only to those who have a theory to prove and look for hidden symbolism in what they suppose to be the true Jungian manner. We are given no first-hand information about the circumstances under which these compositions were in fact written, or the type of teaching adopted in the school which the children attended. Three are clearly reminiscent of poems in children's reading-books, and one is almost a word-for-word reproduction. Our authors do not seem to have questioned the children about the sources from which their 'inspirations' were derived, or to have made any attempt to verify what (so they imagine) this or that child "would seem" to have been feeling, or "must have" experienced in order to produce or reproduce these versified fancies.

We come finally to the problem of education. Here two conflicting voices seem clearly discernible. In the latter half of the book, which I suppose to have been written by Mrs. Branch, the arguments are directed strongly against any kind of separate school specifically for the gifted. Scholarship systems, such as that instituted by the London Education Authority, whereby the brightest 2 per cent. of the school population were enabled to enter an 'established grammar school' or a 'public school' like Christ's Hospital, and the still more ancient system of scholarships to universities, are deprecated as altogether behind the times: "the comprehensive school," it is said, "takes into account the *whole* child in a way that grammar schools and public schools cannot possibly do; within each class the pupils can be divided into very small groups, and because the atmosphere in such a school will be influenced by worldly concerns, there will be an enrichment extending far beyond the purely academic."

On the other hand, the first part of the book takes the opposite view. We are given brief case-histories of ' under-achievers ' whose gifts were either unrecognized or ignored, and who were left in schools where they felt thwarted because they were " mixed up with pupils of far less intelligence than themselves," or " taught by teachers whose abilities were less than their own." We are told how difficult it is for the child of exceptional ability " to expand in a conventional scholastic set-up," and of his " need for a far more specialized and elastic curriculum." We hear of Mary, who " was so dissatisfied by her dull school that she *insisted* on going to a neighbouring grammar school," and of Alan for whom even the teaching at the grammar school (where he was top of his form) failed to come up to his requirements : " the teacher in our biology lesson puts everything *too* simply : I can't understand a thing." As the author adds, " the reason for under-achievement could not be expressed more graphically." An entire chapter—one of the most illuminating in the whole section—is devoted to describing and applauding the merits of " Millfield, one of the few public schools where the child with a really high I.Q. can get a place, whether the parents can afford the fees or not." Indeed, several local authorities, we are told, are now willing to cover the fees, though many who control the purse-strings still seem convinced that such children can be adequately catered for within the ordinary system, and come to no harm by working in a class of forty with a wide range of I.Qs." ' Prospective pupils ' are assessed and selected, not on the basis of the familiar group test, but by a series of individual tests—the Terman-Merrill for intelligence (often supplemented by ' projective tests ') and ' diagnostic tests ' for mathematics and English.

Other chapters contain briefer accounts of several similar ' experimental schools,' including Dartington Hall. One of the most interesting describes an experiment which is being conducted at Brentwood Training College in Essex. Here Dr. Bridges, Senior Lecturer in Education, has established two classes for highly gifted children drawn from four primary schools in the neighbourhood. Mr. Robb, psychologist to the education committee, who is responsible for the testing, has been good enough to send me details, which, I hope, may shortly be published. To many the most surprising result will be the large number of children obtaining I.Qs. from 145 to 170 or over, eleven in the 10-year-old age-group at one school alone. But this is quite in accord with figures that I myself have reported elsewhere (*Brit. J. Statist. Psychol.*, XVI, pp. 175 and refs.) ; the data collected seemed clearly to indicate that current assessments of the number of highly gifted children in our schools (assessments based on a theory of a perfectly normal distribution) grossly underestimate the actual number.

A long chapter is devoted to schools for children with special talents—the Royal Ballet School at Richmond Park, the Yehudi Menuhin School (" where selection is on the strength of sheer musicality "), and the Sevenoaks Public School for the ' technologically gifted.' Unfortunately, in the effort to be popular and picturesque, the descriptions are padded out with a mass of irrelevant detail and lengthy records of chatty journalistic interviews, while the precise particulars as to the mode of selection, curricula, teaching-methods, and results obtained are slurred over as of no general interest.

There is one additional type of special talent school which I should like to suggest—namely, a school for the mathematically gifted. The studies of children with special abilities, carried out from time to time by my colleagues and myself in London, showed that an aptitude for mathematics, like the aptitude for musical ability or ballet-dancing, develops at an unusually early age. Most of the ' infant prodigies ' reported by teachers are children who show an exceptional precocity, interest, and facility in numerical work ; and this is in keeping with biographies of mathematical geniuses, like Pascal, Gauss, Hamilton, Ampère, the Bernoullis, and many others (cf. *Brit. J. statist. Psychol.*, IV, pp. 154 f.). As we are nowadays constantly reminded, the lack of properly trained mathematical specialists is one of the gravest defects in our present educational system ; and of recent years mathematics itself has developed into such a highly complicated subject that it is desirable to begin the specialist's training at the earliest possible age. " How many gifted children have been lost as mathematicians and scientists," Mr. Cash very pertinently asks, " owing to the fact that the teaching in these subjects was so dull ? " But it is not merely the dullness of the teaching, it is rather the level of the teaching, that holds such children back. And after all, it is quite impossible to expect that the ordinary type of classroom teaching should be able to cope with the exceptional needs of a ' potential genius ' in mathematics whose ' mathematical quotient ' is 140 or over. In point of fact, a school such as I have envisaged already exists at Akademogorok in Siberia ; and Professor Lightfoot has recently put forward a persuasive plea for a similar institute in this country.

In conclusion, let me say that, if I have dwelt at undue length on the inaccuracies and misconceptions that have crept into the theoretical portions of the book, that is because they relate to points on which the attention of teachers and educationists has hitherto been chiefly concentrated : as a result the general effect has been to undermine rather than to support the arguments of those who would like to see ampler and better provision made for pupils who are exceptionally gifted. However, these are not the parts which will interest the general reader, for whom after all the book was primarily intended : he will chiefly be interested in the practical possibilities. And there is one all-important fact which the authors have certainly succeeded in bringing out : scattered up and down the country, almost unknown to each other, there are, as they put it, " isolated pockets of enterprising teachers," who have fully recognized the need to discover and cater for children of exceptional gifts and abilities, and have already worked out highly effective methods of achieving these ends.

KALTON, GRAHAM (1966). *The Public Schools : A Factual Survey.* Longmans, pp. xxxii+179, 14s.

By R. R. DALE

Probably the most important part of this book is the section which compares attainment in the public schools with attainment in the state grammar schools. The reviewer considers its statistics to be not quite as impartial as they appear at first sight, and its implications are seriously misleading. The statistics are open to criticism on the following counts :

(1) Nowhere is it stated that the external examinations taken by the independent schools are frequently of a different board from those taken by the majority of state grammar schools. Though the boards attempt to equate their standards, this is not easy to achieve. Moreover, the standard of each board is in part set by adjusting the raw score pass mark to ensure that a fairly steady percentage of candidates will pass each year. This procedure, wise though it may be, must make it rather more difficult for boards to achieve parity of standard.

(2) In the comparison at ordinary level, the author awards an ' O ' level pass to all public school boys who passed in a subject at ' A' level but did not, in fact, sit the ' O ' level examination in that subject. He tries to justify this by saying that in many public schools it is customary to by-pass the ordinary level in the subjects later to be taken at advanced level. None-the-less, the processs seems dubious—the advanced pass is counted both as an advanced and an' O ' level pass, and the possibility of failure at ' O ' level in the subject ruled out. The same concession is not given to pupils in state schools—though it is true that the number of their pupils who by-pass the ' O ' level in the same way is probably small. Also included in the public school totals are ' O ' level passes awarded on the 'A' level examination, but they are excluded from the state grammar school results. In one board's examination in 1966, such passes were often between 10 and 20 per cent. of the number of 'A' level passes awarded, according to the subject. Though these differences in the treatment of the figures are pointed out in the book, the comparison is none-the-less made. The general impression given by the statistics is then quoted by reviewers and others without any mention of the qualifying clauses, to which the author gives inadequate stress.

(3) In the comparison at advanced level the statistics are handled in such a way that the state schools start off with the impossible burden (for fair comparison) of 58 per cent. of pupils ***not taking*** the examination, thereby heavily reducing for these schools the percentages possible in the pass section (cf. Table 6.10). While upholding the comparison as fair, the author sees this point and presents a Table (6 12) which includes only those attempting the examination. In his comment he then selects ' two passes ' as his yardstick ; this happens to be the one which most favours the independent schools. Even here the maintained grammar schools have a slightly higher percentage of pupils with two or more passes than do the two predominant groups of the independent schools (boarding and mainly boarding). If ,however, three or more passes is used as the yardstick, the difference in favour of the maintained grammar school increases appreciably. At age 17, the maintained grammar schools have 53 per cent. of candidates gaining three or more passes, opposed to the 40 per cent. of the above independent schools; at age 18 they have 56 per cent. as opposed to 44 and 43, and at age 19 some 53 per cent. compared with 42 and 46. It should be said that the direct grant schools (apart from those which are mainly boarding) do better still. The reviewer draws no conclusions, but is merely concerned to show that the data can be viewed in a very different light.

(4) For reasons additional to the above, the opposing sets of figures are not necessarily on the same basis. One difference is that the maintained schools' figures reduce to two 'A' level passes all those reported as having passed in three ' mathematics ' or three biological subjects, but there is no mention of this restriction in Kalton's figures. Another is that while the maintained school figures exclude candidates who have passed a second time in a subject, it is not certain that this would always be done in the public school figures. The instructions for completing the forms for ' O ' level results are : " *A subject passed at ' A ' level is also to be included as an ' O ' level pass. The general paper is to be counted as an O-level subject. If the subject is passed more than once, only include it once in the total. For example, a boy who has passed A-level English and O-level History has two O-level passes ; a boy who has passed A-level English and O-level English has only one O-level pass.*" This instruction about counting one pass for a subject passed twice appears here to refer only to subjects taken at both ' O ' and 'A' level, and to apply to the entering up of the O-level results only. The 'A' level instructions make no mention of repeat passes, so presumably they are included in the public school figures and excluded in the others. Moreover, the questionnaire to heads instructed, " If mathematics is passed as a double subject, write Mathematics (2)." Clearly for Pure mathematics heads would write Mathematics (1), but is there not a danger that some of them might class ' Pure and Applied Mathematics ' as a double subject, and write Mathematics (2) ?

(5) For the author to drag in a comparison with " all maintained secondary schools, including secondary modern," is a surprising procedure, which may indicate an unconscious bias. To enter a public school pupils have to pass an examination : to enter a secondary modern school they have to fail one ! In addition, the public schools can discard their failures—or parents can decide that their fees are being wasted—whereas the secondary modern school must perforce keep every one, and admit even pupils who cannot read or write English.

(6) At the public schools the brighter boys often take the ' O ' level examination a year earlier and therefore have an extra year before leaving, in which to accumulate more passes. As, also, candidates tend to by-pass the O-level in their future advanced subjects, they can, therefore, concentrate on a smaller number of subjects, which should improve their pass rate. Though probably not as important as some of the other points it ought to be mentioned.

(7) The figures given for the academic successes of public school pupils who ' failed the 11 plus examination ' are also misleading. Only a small proportion of public (particularly independent) school entrants attended a state primary school—and as most of them were educated privately, their tuition would vary from very good to very poor, and it is, therefore, not surprising that some candidates who

failed the 11 plus examination did better later. Many of their pupils, also, would be aiming primarily at Common Entrance and would take the 11 plus as a side issue. We know also that most of these pupils are of high social class and this is where the academic 'improvers' tend to come from, in any school. In addition, though Kalton places these failures in the secondary modern school and admits they would probably be among the good pupils of that type, he fails to realise that because of the variation in the 11 plus rate to the grammar school from say 40 per cent. to 10 per cent., an *appreciable percentage of his 'failing 11 plus' pupils will, on entry, be academically better than many pupils in grammar schools.*

So much for the statistics themselves, but the implications which derive from them and which escape occasionally in the author's comments are also misleading in other ways. The chief of these implications is that the teaching is better in the public schools than in the maintained grammar schools, as assessed by the standard of attainment reached by the pupils, and it should be said at once that this cannot possibly be proved by the figures presented, mainly because the nature of the entry to a large section of the schools makes the comparison invalid for this purpose. The direct grant schools cream off many of the very best pupils from their own areas, thus depriving the state grammar schools and enriching themselves. The same process occurs, though to a lesser extent, with the independent schools—particularly the day schools. Any comparison of later attainment levels becomes something quite unreal and even laughable if it were not for its influence on an unversed public. We even get a reviewer writing " The survey is a spectacular confirmation of the public schools' claim to deliver the academic goods." (*Times Educational Supplement*) Incidentally, the report itself shows that the direct grant schools have a far better academic attainment level than the independent boarding and 'mainly boarding' schools—as one might expect.

Another reason why the comparison is misleading is the great social class difference in the intake of the two groups of schools. Research has shown repeatedly that pupils from the higher social classes tend to improve at the grammar school while, on the average, those whose fathers are unskilled and semi-skilled are more likely to deteriorate. The public school pupils are predominantly upper social class, while those in maintained grammar schools are predominantly middle and lower class. Again, is there not something bizarre in a comparison between public schools which, because of their wealth, have appreciably smaller classes and can pay higher salaries and thereby attract teachers with slightly better degree qualifications, and state schools who cannot afford these things ? Finally, the book compares those schools which, because of their academic standards, have been admitted to the Headmasters' Conference, with unselected state grammar schools, which scarcely seems fair.

It is said in the Introduction that the public school system is not a closed one, but in fairness it might have been added that, apart from the direct grant schools, the hole through which one can enter is particularly small and open only to the sons of the well-to-do who can offer the Latin which is compulsory for most public schools, and to a small percentage of scholarship holders.

In spite of these criticisms this is a useful book of statistics if read with caution and insight. Unfortunately, the figures will often be taken at their face value and will mislead many educationists as well as the public. Finally, though Mr. Kalton may well have tried to be impartial, he should have stated that he himself was educated at a public school.

BOOK REVIEWS

ARCHAMBAULT, R. D. (Edit.) (1965). *Philosophical Analysis and Education.* London: Routledge and Kegan Paul, pp. xii+212, 28s.

Most departments of education and training colleges offer their students courses entitled " The Philosophy of Education " or " Educational Theory " or, even more grandiosely, " The Meaning and Purpose of Education." But many who have suffered such courses may be inclined to agree with a remark made by one of the contributors to this book: " Most of what is written on the ' philosophy of education ' is monumentally dull (and) seems utterly divorced from reality " (Wilson, p. 158). He might also have added that such writings appear equally divorced from much that has recently been accomplished by philosophers.

This book represents a commendable attempt to change all this. It is a collection of essays, all written expressly for this volume, by nine British philosophers, and which are preceded by a short but useful introduction by an American editor. In general, their purpose is to show the kind of contribution that can be made by the modern, analytical philosopher towards unravelling the complex conceptual issues that arise whenever men think seriously about traditional educational problems. True, the first two essays, by L. A. Reid and E. Best, present rather different conceptions of the philosopher's task, but the remaining authors tend to follow Best's recommendations in making modest and limited analytical contributions towards solving particular educational puzzles. Thus, L. R. Perry, R. S. Peters and P. H. Hirst all consider, in their different ways, what is meant by ' education '; J. P. Corbett, J. Wilson and R. F. Atkinson, in a further section, concern themselves with the concept of ' teaching '; and finally, A. P. Griffiths analyses the notion of ' university.'

The result is a book which, as a whole, will not easily be understood or appreciated by those lacking a background in modern philosophy. This is not to deny the quality, clarity—even *panache*—of many particular contributions: the papers by Best, Peters, Corbett and Wilson, for example, are well worth reading. But the reviewer's feeling is that most students, unhappily struggling through their courses on the ' philosophy of education,' would gain a much clearer and quicker understanding of the modern philosopher's contribution to education from D. J. O'Connor's excellent *An Introduction to the Philosophy of Education* (1957), than from the present volume.

G. W. PILKINGTON.

BRITTON, J. N., MARTIN, N. C., and ROSEN, H. (1966). *Multiple Marking of English Compositions: An Account of an Experiment.* Schools Council Examinations Bulletin No. 12. London: H.M.S.O., pp. vi+50, 3s.

Pamphlets reporting research, if they make an important contribution, tend to go out of print quickly. This Bulletin is, therefore, a wise investment at three shillings a copy. Thorough in design and execution, clearly and concisely reported, the experiment could be used to advantage as a model of research in educational measurement.

Five-hundred G.C.E. scripts were assessed independently by three teams, each of three markers working rapidly and judging by their general impression of the piece of writing, together with a fourth marker who applied a code of penalties for mechanical errors. These combined marks were compared with the G.C.E. marking and with two other markings in the traditional style. In addition, a series of ten compositions by the candidates on different types of subject, written at monthly intervals and each marked by twenty examiners, provided a broadly based criterion of writing ability. Teachers' assessments were also obtained. In all, thirty-five different markings are brought together.

Those familiar with previous work on the marking of compositions will know that the subject presents unusual problems in defining reliability and validity. The intricate design of this study is necessary to review the many aspects involved.

Multiple marking in fact, proved superior. (One of the reliability coefficients on page 25, given as ·18, is a misprint for ·81.) Those who have for long depended on Hartog and Rhodes now have a more adequate report available.

The Department of Education and Science made a grant of £1,000 towards the cost of this investigation. The estimated cost of marking a G.C E. script by the traditional method is 3s. 5d. Reckon the numbers involved in any one year, and you have an answer to those who are still inclined to question the value of expenditure on educational research. JOHN NISBET.

DIENES, Z. P., and JEEVES, M. A. (1956). *Thinking in Structures.* London: Hutchinson Educational, pp. 126, 21s.

The authors describe their efforts to study how children and adults sort out the seeming chaos of their environment into some kind of order, and so think in structures or patterns of relationships. Thought structures based on mathematical groups of two and four elements were involved in the experiments. The subject had to observe a card in a ' window ', play one of his own cards (selection situation) or have a card played for him by the experimenter (reception situation), and then forecast what the next card appearing at the window would be. Standard questions were asked after both the two-game and the four-game, including one which requested the subject to say how he thought he had succeeded in solving the problem. It was possible to classify the evaluations to this question into operational, pattern and memory types, or a combination of these.

The placing of the evaluations into this hierarchy seemed to be valid. For subjects in the selection situation there was a close relationship between the hierarchy and the number of instances necessary to solve the problem, while for testees in the reception situation there was close agreement between the hierarchy and the total number of errors and other scores. There were also some interesting sex differences in performance. Further, the results suggest that it might be better to begin with a more complex task and continue with a simpler one from the point of view of explicit evaluation. Curious though is the fact that while subjects tended to give the same kind of evaluation to both the two-and four-games, there was little correlation between the ability to solve the two tasks.

Dienes and Jeeves have opened up new ways of studying thought processes and they have brought out the possible educational implications of their study. At the same time, it must be pointed out that we do not know if children make order of these somewhat artificial tasks. However, readers of this *Journal* interested in cognition, will enjoy this work and they will look forward to an account of the more recent experiments carried out by the authors using tasks embodying groups of up to nine elements ; although it is hoped that such an account will give more details about the subjects tested than does the present Monograph in respect of the subjects involved in the present study. It is a pity this Monograph does not have an index.

K. LOVELL.

EBEL, ROBERT, L. (1965). *Measuring Educational Achievement.* New Jersey: Prentice-Hall, pp. xii+481, 64s.

The book is primarily concerned with the varieties of educational tests which can be prepared by teachers and lecturers for use in their own classes. It deliberately excludes direct discussions of standardised attainment tests or of other types such as tests of aptitude or personality. In the reviewer's opinion it was a wise decision not to attempt to cover the whole field of educational tests with the inevitable risk of superficiality.

The author states in the preface that the book is intended for practicing teachers who have had no previous special training in educational measurement. He has tried to steer a middle course between excessive simplicity and excessive thoroughness. In spite of this declared aim, however, he has not shirked the discussion of difficult topics such as a derivation of the Kuder-Richardson formulae 20 and 21 for determining reliability from a single administration of a test. As the author admits, the book may sometimes tax the student's powers of comprehension. While the style is clear and straightforward, students will require to have some previous background knowledge to read some of the later chapters without considerable effort. The book would scarcely be suitable as a basic text for the average class in educational measurement in a college of education. It would, perhaps, be more suitable for students taking courses for the M.Ed. degree or for a Diploma in Education, and it would be an excellent book of reference for lecturers teaching such courses. The treatment is essentially practical. The book gives numerous examples of the procedures involved in preparing test-items, illustrating many of the commoner types of faults with a wealth of sample items in different fields of study.

There is an excellent chapter on how to improve test quality through item analysis. There is also a chapter consisting of some 40 pages on basic statistical techniques, but this is scarcely adequate to provide the background for understanding the more technical discussions in later sections of the book. The treatment of reliability is lucid and cogent, but the chapter on validity is perhaps less satisfactory. About ten different types of validity are discussed but since these are not all distinctly different from one another the general effect tends to be less clear-cut.

Among the good features of the book, the following might be mentioned: There is a concise summary of important points and a well-chosen list of references for each chapter. There are numerous tables summarising relevant statistical data and there is a very substantial glossary of technical terms at the end of the book. There are excellent practical recommendations on aspects such as how to judge the quality of a class-room test or how to construct a specific marking system.

The book can be recommended to all teachers and research workers who require information on how to construct their own achievement tests.

I. MACFARLANE SMITH.

EIFERMANN, RIVKA R. (1966). *Scripta Hierosolymitana, Vol. XIV, Studies in Psychology*. London: Magnes Press, Oxford University Press, pp. 230, 43s.

This volume comprises ten papers divided into four sections and written by some fifteen authors. Perhaps the most impressive feature of the work is that it comes from a department opened only seven years ago, and as such represents a not inconsiderable feat. Kugelmass, the acting head of the department, in his foreword to the volume, offers as one of his aims the opportunity the book provides for the department to be viewed as a whole. This it does, and although there are some fields notable by their absence, it shows quite a breadth of interest.

In the first section of the book, Studies in Development, there are three papers. Gewirtz has contributed a fifty-page article dealing with his researches on the development of smiling in different groups of Israeli infants. This is a detailed and important paper for anyone interested in the development of behaviour in infants. Kohen-Raz on movement representations and their relation to conceptual thought is much more difficult to understand. Many of the concepts he uses are inadequately defined and the paper is made difficult to read because of the large number of footnotes. Kugelmass and his collaborates describe a study in the development of moral judgment in the final paper of the section.

Three papers make up the section on social behaviour and personality; one on the development of the superego, one culture and game strategy and an unusual and interesting study on the reactions in Israel to the Eichmann trial. The third section of the book contains what is, perhaps, the most important paper, "A faceted definition of intelligence" by Guttman. This is an introductory paper which outlines what may well be an important development in the measurement of intelligence. Ortar's paper on entrance examinations for the department of psychology of the Hebrew University is likely to be of interest to all who are concerned with the problem of selecting students to read for degrees in psychology. Ortar and his colleagues have constructed a set of intelligence and achievement tests as a means of selecting students. On the whole, the data from the follow-up study which has been conducted indicate that the sets have ' considerable ' validity. The work is continuing and some of the tests remain to be validated.

The fourth and final section contains two papers on what are collectively described as " Cognitive Processes." Both of these, one on selection strategies in concept attainment the other on redundancy in a card sorting task,, are short and Schlesinger in his paper on redundancy reports the details of an experiment. This last paper would certainly have been more appropriately published in one of the experimental journals.

Inevitably, in a book of this nature, the contributions are somewhat uneven and one is left with the feeling that much of the work would have been more appropriately published elsewhere. There is also a certain parochialism about many of the papers but perhaps this is to be expected. Despite these criticisms, however, the volume marks a substantial achievement in seven years. The next seven years may see an estension of the department's interests into comparative, physiological and experimental psychology. The reader is invited to consider the volume as an overview and it would seem to be these fields which are lacking. KEVIN CONNOLLY.

HILDRETH, GERTRUDE H. (1966). *Introduction to the Gifted.* New York: McGraw-Hill, Inc., pp. xvii+572, 64s.

The name Novosibirsk nowhere appears in this all-but-memorial compendium of information concerning the psychological, social, and educational problems that appear if we recognise the existence of a minority of individuals whose ability, on the one hand, to learn and, on the other, to contribute to the wisdom and the welfare of mankind, is significantly higher than that of the generality. This reviewer, at least, refuses to believe that an indefatigable propagandist like Hildreth would have thrown away this master-card in the interest of a cheap psychological chauvinism;

As Hildreth herself points out, this book has grown out of her own personal experience with gifted children in public and private schools, dating back over forty years. Beyond that, however, it represents a massive stock-taking of the hundred years of research and propaganda on the subject from Galton through McKeen, Cattell, Spearman, Terman, and Burt to Guilford, Wolfle, and Flanagan. In 1966, when the attitude to the determinants of individual differences in ability is, in this country, uncomfortably equivocal, and when emphasis upon intellectual or academic excellence has become almost a psychological indecency, Hildreth's openly avowed philosophical commitment is not likely to be immediately popular. It is in our best interest, Hildreth argues, " to abandon the fiction that recognising and fostering exceptional gifts and talents is undemocratic or that special school programs for talented youth are unfair to those not include ." We may be ready to accept that " the most valuable investment any society can make is the development of the full potential of its human resources," but Mr. Crossland's political prescription for how that is to be done is liable to be somewhat different from that of one who is convinced, as Hildreth is, of the genetic origins of giftedness. While she would like environments and method of instruction to be ' superior,' it is, nevertheless, her belief that exceptional results will be achieved only if the learner is both ' bright ' and ' interested ' before we start.

Harding has noted as a characteristic of our culture that " most people's sentiments are organised in a way that makes it easier to lavish attention on the unfortunate than to help the gifted to make the most of themselves." The eighteen chapters of Hildreth's vast ' review of the literature ' is directed at redressing the balance, and it is fair to notice that the writer's survey is not limited to American work or to the American scene. The book is not difficult reading, and inevitably it is often superficial and sometimes naive, if not, indeed, sentimental. But its courage of the relevant psychological investigations and educational practice (the latter chiefly in the United States) is very extensive. Most of the significant issues are touched upon, even if none is fully or penetratingly examined—the nature of giftedness, the criteria and methods to be employed in the discovery and selection (*sic*) of the gifted, the problem of ' creativity,' the kind of educational provision that can (or should) be made for students of high ability at all states from primary school to university, and even—an astonishingly weak section, this—the kind of teachers they need. From the point of view of the psychologist, the discussion is generally limited and unsatisfying, but for the teacher in training—and for those who teach him—the material that the book brings together and the point of view that it asserts merit serious consideration. A ' message ' is not necessarily simple-minded because it is simply put. What C. F. Carter has called " the economic use of brains ," is one of the crucial problems of the contemporary world. It is fairly certain that at the educational level, the comprehensive nostrum is not the answer to it.

GEORGE SETH.

KELLMER-PRINGLE, M. L. (1965). *Deprivation and Education.* London: Longmans, pp. vi+311, 42s. 6d.

A book title which implies that the contents merit the consideration of educationists is frequently misleading, but this cirticism cannot be applied to this work by Dr. Kellmer Pringle. The book is obviously written for the reader with a professional interest and knowledge, and will not, therefore, achieve the readership figures of one with wider public appeal. This is unfortunate, as there is a message to sell, if not a book for its own sake.

One drawback which is not immediately apparent to the reader, is that the book is effectively in two parts. The first part is a collection of research studies carried out by Dr. Kellmer Pringle and some of her former students, on the effect on intellectual, emotional and laguage development, of deprivation such as found among children in residential care, or rejected by parents, or in a culturally and emotionally impoverished environment. The second part provides a more detailed discussion of earlier research findings combined with the results of wide knowledge and experience of the subject matter, which includes the educational and emotional needs of deprived children, the roles of adults, such as nurses or child care workers, in whose hands such children may be placed for short or long-term care, and the needs of the family—all of which, as Dr. Kellmer Pringle suggests, may provide a guide to early prevention of recurring family problems. She also puts forward the proposal, as did the Kilbrandon Report, that there be an integration of the various bodies concerned with the welfare of children and a comprehensive service provided for all parents [illegible]ng family difficulties.

The early chapters provide an admirable model of reporting research studies but the meticulous detail is a distraction from the continuity of presentation of the findings, which have a significance for students in the fields of education and child care. Inevitably in summarising the results, and in making cross references,there is a degree of repeition which becomes irritating.

The later chapters of the book are more rewarding to the reader who has been attracted in the first instance by the title and the fly leaf. Here, Dr. Kellmer Pringle presents cogently information and advice on the current position as regards work with children who are deprived socially, emotionally and educationally, and as she rightly stresses, this is a neglected field of study which may only now begin to receive the sympathetic attention of officialdom, if the values of early preventive measures can be shown to bear fruit.

As a work which is based on practical experience and sound research, this book is to be commended as a useful source for discussion, information and reference to serious students and workers in the field of child care. There is, however, its additional value, expecially in the later chapters, to students and teachers more concerned generally with child development and the broader application of education to all children, and to those who may eventually be asked to implement such recommendations as those proposed by Kilbrandon and Plowden.

RICHARD L. PEDDIE.

MIALARET, G. (1966). *The Psychology of the Use of Audio-Visual Aids in Primary Education.* London: George G. Harrap, pp. 225, 30s.

This book is " intended primarily for teachers of psychology and education in teacher training colleges and for research workers."

It is not, as its title might suggest, a comprehensive text setting forth the principles underlying the use of classroom aids. It is, in fact, an expanded report of a UNESCO seminar held at Caen in 1962. Contributions by seventeen participants have been linked together by the author who is Director of the Laboratoire de Psychopédagogie at the University of Caen.

In the first half of the book an attempt is made at a general analysis of the perceptual components of the audio-visual teaching situation in terms of the physical characteristics of stimuli, the prevailing conditions of illumination, the nature and ' information density ' of the message to be communicated and personality variables in the pupil. There follows a detailed treatment of selected aspects of visual and auditory perception (such as size, movement and speech recognition) and the author attempts to formulate some of the problems posed by simultaneous reception of visual and aural information. In the second half of the book the author selects one particular topic in emotion—that of empathy in young film audiences—which he analyses by reference to several research studies. He then deals with one cognitive problem—the comprehension of still and moving film and finally with the use of film in the creation of motivation. The book concludes with a statement of the author's views on the part that audio-visual aids should play in teacher-training.

As the author readily admits, it is hardly possible to do more than discuss a small selection of experimental findings. Many of the enquiries reported are by French psychologists and may not be widely known in this country. Some are of great interest—such as the study of the reactions of children of various ages to a Chaplin film. Whether fear or joy was evoked by any particular incident in the film seemed to depend very clearly on the age and sex of the viewer. A careful study by Zankow in the U.S.S.R. showed the superiority of teaching leaf-cell structure to botanists by visual rather than by descriptive methods. The examples used are good ones, but the argument is often ponderous and the discourse fragmented. The book never remotely approaches the lofty promise of its title. One looks in vain for any evidence of an attempt to relate the discussion or illustrative material to modern theoretical concepts—such as those of learning theory—or indeed, to any systematic framework. The book seems to have suffered in translation. There is, however, a lack of crispness in the exposition which goes beyond problems of translation; the ratio of words to ideas is often much too high. There is more than a hint of name-dropping and jargon is shamelessly used.

There is a good deal of speculation enmeshed in the author's arguments—but the speculation is often of great interest. Indeed, one of the merits of the book for English readers may well be its uninhibited assurance and its freedom from many of the current methodological preoccupations of Anglo-American psychologists. Perhaps it is unfair to look for comprehensiveness of coverage and precision of presentation in a book which is trying to break so much new and worthwhile ground. It will certainly provoke useful discussion among teachers and students—though it may give them little practical help in the classroom. Theoreticians will be infuriated by its dogmatic pronouncements and clumsy exposition, but research workers will find much here of interest—not least in its useful bibliography. JOHN LIGGETT.

NIKELLY, A. G. (1966). *Mental Health for Students: A Guide for Adjusting to College.* Illinois: Charles C. Thomas, pp. xv+208, $7.50.

The objective of this book is " to clarify mental health problems for the student and to explain to him the various reasons that keep students from performing adequately." Nikelly emphasises to the student reader (p. 21) that " knowledge alone about human behaviour will not change the student or cure his problems," but sees his book as giving " some facts about the concept of normal behaviour and . . . what the student can expect from certain behaviour problems in the university setting," thereby enabling the student " to cope with his difficulties in college."

There are seven chapters. The first, a general account of the nature of mental health, particularly in terms of self acceptance, is followed by one on the needs and feelings of the student and by one on the special areas of maladjustment in college. There is an extensive survey of the mechanisms of ego defence and other protective behaviours which are seen on the whole as benign techniques which the insightful student can accept as preserving his integrity. One chapter outlines the characteristic symptoms found among disturbed college students, another the main types of serious maladjustments. Finally, there is an account of the people who can help a disturbed student, how they do their jobs, how they can be found and what to expect of them.

At first sight, this would seem a hypochondriac's *vademecum*. It could be : but it is also, on the whole, a good attempt at achieving the author's aim. It could help students to gain useful insights and so facilitate their solution of the developmental tasks facing them and their tolerance of the stress of their situation. The author walks the razor edge of giving the student a reassuring perspective on the variations in ' normal ' behaviour and information on the background to such variations, while maintaining the implication (helpful to the general run of students) that they can take charge of their own lives and manage their growing up. Only occasional comments (" Life would be uninteresting if all of the problems of living were suddenly solved ") introduce a moralising tone that might turn students away.

The problem of presenting technically complex concepts in lay terms does lead, however, in Chapter 1, to a generality of comment that lacks the particulate immediacy (such as one finds in Truscott's *First Year at University*) that would seem useful in communicating to the student reader that the book has something of value to him. The non-technical language presents no barrier ; but a terser style and even more use of chapter sub-headings would recommend it to the reader seeking assistance.

At times, the book might be addressed to those who will deal with students as much as to students directly ; and, in Britain, it might indeed be of most use in. introducing university and college advisers of studies and tutors to the background in psychology of the problems posed to them by their students. T. E. CARRUTHERS

PAGE, E. B. (1964). *Readings for Educational Psychology*. New York : Harcourt, Brace and World Inc., pp. xii+404, $3.95.

Although this collection of Readings is designed to accompany Cronbach's *Educational Psychology*, 2nd Edition (1963), it is rightly claimed by the compiler that it could supplement any educational psychology textbook. Included in the thirty-seven papers reproduced are contributions from Skinner (The Science of Learning and the Art of Teaching), Jackson and Getzels (Dissatisfaction with School among Adolescents), Bandura (Social Learning Through Imitation), Gagne (Training and Principles of Learning), Bruner (Readiness for Learning), Humphreys (Transfer of Training in General Education), Bronfenbrenner (Soviet Methods of Character Education), Becker (Social Class Variations in the Teacher-Pupil Relationship), Frandsen (Learning: An Integration of Theories), Mussen and Jones (Late and Early-Maturing Boys), Symonds (Classroom Discipline) and Janis and King (The Influence of Role Playing on Opinion Change). This list provides some measure of the range of topics and the standing of the contributors.

It is, however, disappointing to find a number of dated contributions. Page remarks in the Preface, " Educational Psychology is a fast-changing field in a century of change." It is, therefore, a little surprising to find that of the thirty-seven readings, twenty-one are pre-1960. Two are from the 1930's and three from the 1940's. Of these five, only Symond's article would appear to justify its inclusion.

The Editor has made a commendable endeavour to help students evaluate the papers by posing in the introduction to each contribution a number of questions directed mainly at the interpretation of experimental results by the writer. This is a helpful feature, which should encourage discussion where this book is used for tutorial or seminar purposes. The inclusion for example, of Flesch's article " Phonics *vs.* no Phonics " as the first reading immediately highlights the questionable interpretation of research results. The balance, however, is redressed by the second reading (Carroll: The case of Dr. Flesch) which is well documented with ample references. The same cannot be said of Thelen's article in which the case against programmed materials is stated.

On the whole, a valuable collection of extracts. Many contain vivid classroom illustrations which would make the book worthy of inclusion in College of Education Libraries. A. MILNE.

SCHEIN, E. H., and BENNIS, W. G. (1965). *Personal and Organizational Change through Group Methods: The Laboratory Approach.* New York: John Wiley, pp. x+376, 62s.

The present volume is a very difficult book to evaluate. It is concerned with ' laboratory training ' which, for the uninitiated, refers to attempts to change attitudes to self, to others and to questions involving human relations largely by means of free ' unstructured ' small-group meetings (' T ' groups). These methods, which, the book informs us, have been in existence for about 18 years, appear to have spread quite widely in the United States, and more recently to some extent, in Western Europe. The approach has clearly been much influenced by Lewin, and some of the language used is reminiscent of him. It also has much in common with group psychotherapy, though as the authors indicate, this similarity may be greater or less according to the particular view of ' laboratory training ' which is taken by different practitioners. The courses are sometimes residential, sometimes not, sometimes part-time and sometimes full-time and may be highly concentrated or spread over a longer period of time. Some of the comments made by people who have taken these courses indicate that their impact has been considerable. It is, however, a very difficult field in which to validate the procedures used, to determine how much effect they have, how enduring it is, and precisely what it is that is having the effect. It may well be that there are interesting and important possibilities in this kind of thing. A great deal more, however, has still to be done by way of clarification and validation if it is to be accepted. There are also, as the authors agree, ethical problems concerning the possibilities of ' human manipulation.'

The present book may be of interest to those who already know a bit about ' laboratory training,' but if the present reviewer's reaction is at all characteristic, it is not the ideal introduction for those who do not. It is not sufficiently clear *exactly what* goes on at laboratory training sessions, or *what* theoretical concepts are being used. There is, however, both a selected bibliography and a list of references for those who want to find out. D. GRAHAM.

SEARS, R. R., RAU, L., and ALPERT, R. (1966). *Identification and Child Rearing.* London: Tavistock Publications, pp. xiii+383, 63s.

In psychoanalytic theory the process of identification is a key concept which is held to account—among other things—for the way in which children come to adopt appropriate sex roles, learn to sublimate aggressive impulses, and restrain or inhibit many other forms of anti-social conduct. Freud, however, offered more than one suggestion about the conditions under which identification was assumed to take place. According to the theory of ' analytic identification,' the young child comes to identify with his parents because they are loving and nurturant yet cannot always give him their continuous and undivided attention; the fear of loss of love is all-important. In the theory of ' defensive identification ' the child conceptualizes the parent as a power figure with whom he has an ambivalent love-fear-hate relationship. The child resolves his unbearable yet inescapable conflict by identifying with, and becoming like, the all-powerful parent. The picture is, of course, complicated by the fact that both mother and father can, and do, act as disciplinarians ; and the outcome for boys is assumed to be different from that for girls. This, then, is the psychoanalytic peg from which—in his usual style—Sears hangs his own neo-behaviourist theoretical formulations. The questions he is asking thus have a somewhat mixed ancestry, but they are nonetheless important questions for any theory of child socialization. Why, for instance, do some girls become adventurous and aggressive tomboys while some boys remain timid, shy and clinging ? Why do some children develop an exaggerated sense of right and wrong at an age when other children apparently remain cheerfully amoral ?

This book is a research report of an attempt at a massive frontal attack on numerous problems of this sort. The methods used to collect information about the children studied are both ingenious and varied. Both parents were interviewed separately, using questionnaires covering almost every aspect of child rearing, and mothers were also asked to rate their own children in various ways. The children themselves were observed over considerable periods of time, both in spontaneous play and in various artificially contrived situations either with their mothers or with other children. All the children took part in standardised doll-play sessions, and were also used as subjects in comparable experimental settings. One memorable experiment, in which the child is asked by the experimenter to keep an eye on a living pet hamster which is made to vanish—by a subtle mechanical device—the moment he takes his eyes off it, will no doubt be used to break the monotony in many a dull lecture yet to be prepared. In all, no less than 370 separate ratings were painstakingly made for each child, and wherever possible careful checks for reliability and consistency were built into the assessments. In terms of technical competence the research methods described by this team are magnificant and could well serve as a model to indicate how research in this field should ideally be conducted ; and yet . . .

The end result of all this work is both theoretically disappointing and generally confusing. In part, no doubt, this is due to the complexity of the subject, and to the inadequacy of current theories. In this particular investigation, however, the size and constitution of the sample and the method of analysis used must also have contributed to the general inconclusiveness of the results. The children were twenty-one boys and nineteen girls, between 4 and 5 years of age, who were attending the Stanford Village University Nursery School. The correlation coefficient is used routinely as a crude index of association between any two variables, and boys and girls are dealt with separately throughout. With this small number of subjects an r of 0·43 is necessary merely to indicate the possibility of some non-chance association between any two variables ; and if 100 such correlations are examined we would expect to find approximately five of them exceeding this value fortuitously. To be fair to the authors, it must be added that they are aware of this problem ; and they also find that the proportion of correlation coefficients greater than 0·43 is on average a little higher than 5 per cent. There remains, however, the difficulty of judging which of their correlations should be taken seriously as meaningful research findings. In fact, of the thousands of cross-correlations cited in this volume, only a small percentage are significant at the 0·01 level, and only a tiny handful are as large as 0·7. In these circumstances, all the authors can do is to look for broad directional trends among tables of correlation coefficients ; and this makes the business of interpretation more of an art than a science. Those who (including myself) enjoy the Searsian brand of speculation will no doubt find the book a very useful stimulus ; others may prefer to wait until his theoretical ideas can be put to some more direct and searching test.

JOHN NEWSON.

SHAW, OTTO, L. (1965). *Maladjusted Boys.* London : George Allen and Unwin, pp. vii+168, 28s.

Amid the din of shouting about the treatment of disturbed youth, there is a danger that Shaw's book will be undeservedly ignored. Shaw has created a successful residential ' school ' in which education through to grammar-school level, shared responsiblity and psychotherapy are inextracably interwoven. And this unique microcosm *works.*

The difficulties are not coyly shunned. Nor are suspiciously promising successes claimed. Shaw, quietly, pragmatically, honestly and modestly describes the workings of a ' school,' which I suspect depends largely upon his charismatic personality, sensitivity and ability.

The possibilities for research are tantalisingly enormous. What child is best suited to this kind of situation ? What adults make the best teacher-therapists ? What is the nature of the informal social organization within the formal structure ? How do successful/unsuccessful Former-pupils view the school ? How far could one apply Shaw's methods to children of lower intelligence ?

Teachers and educational psychologists would probably welcome more about the dimly-perceived relationships that develop between pupils and between pupils and staff even in a conventional educational system. L. BLOOM.

NEW JOURNAL

The Journal of Aesthetic Education. Editor Ralph A. Smith, Office of the Superintendent of Public Instruction, Springfield, Illinois.

We welcome the arrival of the first issue of a new Journal in a much neglected field. Its purpose is to " stimulate understanding of the problems of aesthetic education," where aesthetic education is defined as (*a*) " the study and appreciation of literature, music, the visual arts, dance, theatre, regional planning, architecture, and film and (*b*) the aesthetic dimension of all life and education." The editor is the Assistant Professor of Art and Education at Illinois.

The inaugural issue has appeared, but this does not represent the form of future numbers for this one contains only reprints of papers which have appeared elsewhere and we understand that subsequent numbers will be devoted almost entirely to original papers. Among the papers reprinted in the inaugural issue is one from this *Journal* ; Cyril Burt's critical notice of C. W. Valentine's *Experimental Psychology of Beauty.* In the field we could certainly have offered nothing better. Most of the other papers are less specifically psychological, but there is one on *Shaping the Emotions* by Donald Arnstine, which will greatly interest readers in colleges of education who are continually taxed by the problem of the education of the emotions.

The subscription rate of the new Journal has not yet been fixed and for the present it will only appear twice a year, but it is hoped shortly to publish regularly each quarter. Papers should be sent to the Editor at 118, Fine Arts Building, University of Illinois, Champaign, Illinois 61820.

PUBLICATIONS RECEIVED

W. Dennis (1966). *Group Values Through Children's Drawings.* New York: John Wiley, Inc., pp. xiii+211, 53s.

G. A. Ferguson (1966). *Statistical Analysis in Psychology and Education.* New York: McGraw-Hill, pp. x+446, 64s.

Gaston Mialaret (1966). *L'Apprentissage de la Lecture.* Paris: Presses Universitaires de France, pp. viii+121, no price.

Gaston Mialaret (1966). *Education Nouvelle et Monde Moderne.* Paris: Presses Universitaires de France, pp. 169, no price.

D. H. Stott (1966). *Studies of Troublesome Children.* London: Tavistock Publications, Ltd., pp. xi+208, 32s.

M. Dunfee and H. Sagl (1966). *Social Studies Through Problem Solving: A Challenge to Elementary School Teachers.* New York: Holt, Rinehart and Winston, Inc., pp. xii+386, $7.25.

L. Hollowell (1966). *A Book of Children's Literature.* New York: Holt, Rinehart and Winston, Inc., pp. xi+580, no price.

H. C. Lindgren, D. Byrne and L. Petrinovich (1966). *Psychology: An Introduction to a Behavioral Science.* New York: John Wiley and Sons, pp. xv+560, 60s.

M. E. Hebron (1966). *Motivated Learning: A Developmental Study from Birth to the Senium.* London: Methuen and Co., pp. 264, 35s.

H. J. Eysenck (1966). *Check your Own I.Q.* Penguin Books, Ltd., pp. 190, no price.

J. Bieri (1966). *Clinical and Social Judgment: The Discrimination of Behavioral Information.* New York: John Wiley and Sons, pp. xiv+271, 60s.

John Jung (1966). *Contemporary Psychology Experiments: Adaptations for Laboratory.* New York: John Wiley and Sons, pp. xiii+157, no price.

L. Uhr (1966). *Pattern Recognition.* New York: John Wiley and Sons, pp. ix+393, 68s.

R. S. Illingworth (1966). *The Development of the Infant and Young Child: Normal and Abnormal.* Edinburgh: E. and S. Livingstone, Ltd., pp. ix+378, 37s. 6d.

Wm. W. Wattenberg (1966). *Social Deviancy Among Youth.* U.S.A.: The National Society for the Study of Education, pp. xi+434+vi, 41s.

M. Rousselet (1966). *L'enseignement du Latin.* Paris: Presses Universitaires de France, pp. 150, no price.

Walter M. Lifton (1966). *Working with Groups: Group Process and Individual Growth.* New York: John Wiley and Sons, pp. xii+288, 53s.

Gayle Janowitz (1965). *Helping Hands: Volunteer Work in Education.* Chicago: The University of Chicago Press, pp. 125, $3.95.

John I. Goodlad (1966). *The Changing American School: N.S.S.E. 65th Yearbook, Part II.* Chicago: The University of Chicago Press, pp. xi+319+XCIX, 37s. 6d.

Irwin G. Sarason (1966). *Personality: An Objective Approach.* New York: John Wiley, pp. xvi+670, 65s.

Arthur G. Wirth (1966). *John Dewey as Educator*: His Design for Work in Education (1894-1904). New York: John Wiley, pp. xiii+322, 49s.

I. N. Berlin, and S. A. Szurek (1966). *Learning and Its Disorders.* California: Science and Behavior Books, Inc., pp. xviii+295, $5.95.

British Postgraduate Medical Federation (1966). *The Scientific Basis of Medicine Annual Reviews.* University of London: The Athlone Press, pp. xi+352, 40s.

K. R. Hammond (1966). *The Psychology of Egon Brunswick.* London: Holt, Rinehart and Winston, pp. x+549, 96s.

Edwin Fenton (1966). *Teaching the New Social Studies in Secondary Schools: An Inductive Approach.* New York: Holt, Rinehart and Winston, Inc., pp. xviii+526, $7.95.

Ralph N. Haber (1966). *Current Research in Motivation.* New York: Holt, Rinehart, and Winston, Inc., pp. xii+800, no price.

Reginald D. Archambault (1966). *Philsophical Analysis and Education.* New York: Routledge and Kegan Paul, pp. xii+212, 28s.

J. E. McGrath and I. Altman (1966). *Small Group Research: A Synthesis and Critique of the Field.* New York: Holt, Rinehart and Winston, pp. ix+601, $12.50.

B. Kleinmuntz (1966). *Problem Solving: Research, Method and Theory.* New York: John Wiley, pp. ix+406, 53s.

J. Tiffin and E. J. McCormick (1966). *Industrial Psychology.* London: George Allen and Unwin, pp. xiv+682, 60s.

R. S. Peters (1966). *Ethics and Education.* London: George Allen and Unwin, pp. 333, 40s.

E. Rowe (1966). *Failure in School: Aspects of the Problem in Hong Kong.* Pub. No 3. Hong Kong University Press, Oxford University Press, Hong Kong Council for Educational Research, pp. xvii+167, 22s. 6d.

T. R. Miles (1966). *Eliminating the Unconscious.* Oxford: Pergamon Press, pp. xviii+171, 17s 6d.

B. Babington Smith (1966). *Laboratory Experience in Psychology: A First Term's Work.* Oxford: Pergamon Press, pp. xxiv+240, 25s.

Edward L. Wike (1966). *Secondary Reinforcement: Selected Experiments.* New York: Harper and Row, pp. xii+516, no price.

R. A. Baker (1966). *Psychology in the Wry.* New Jersey: D. Van Nostrand Co., pp. ix+170, 14s.

E. Stones (1966). *An Introduction to Educational Psychology.* University Paperbacks London: Methuen, pp. 424, 18s.

J. Haggerty (1966). *Please, Miss, Can I Play God.* London: Methuen, pp. 149, 21s.

J. E. Sadler (1966). *J. A. Comenius and the Concept of Universal Education.* London: George Allen and Unwin, Ltd., pp. 318, 35s.

D. Morris (1966). *The Biology of Art.* London: Methuen and Co., pp. 176, 42s.

J. H. S. Bossard and E. S. Boll (1966). *The Sociology of Child Development.* 4th Ed. New York: Harper and Row, pp. x+566, 32s.

J. D. Nisbet and N. J. Entwistle (1966). *Age of Transfer to Secondary Education.* London. University of London Press, Ltd., pp. 111, 12s. 6d.

W. Schramm, J Lyle, and E. B. Parker (1966). *Television in the Lives of our Children.* California: Stanford University Press. London: Oxford University Press, pp. iv+324, 22s. 6d.

G. F. Kneller (1966). *Logic and Language of Education.* New York: John Wiley, pp. viii+242, 45s.

R. Stevens (1966). *An Approach to Literature.* London: Longmans, pp. 297, 13s. 6d.

B. J. Elliott (1966). *Modern Times: Hitler and Germany.* London: Longmans, pp. viii+176, 8s. 6d.

P. Groves and L. Stratta (1966). *Lost in the Fog.* London: Longmans, pp. 47, 6s.

P. Ashlock and A. Stephen (1966). *Educational Therapy in the Elementary School.* U.S.A.: Charles C. Thomas, pp. xv+102, $6.75.

L. Blank (1966). *Psychological Evaluation in Psychotherapy.* London: George Allen and Unwin, pp. xi+364, 70s.

J. Cohen (1966). *A New Introduction to Psychology.* London: George Allen and Unwin, pp. 220, 25s.

M. Roff, W. Mink and G. Hinrichs (1966). *Developmental Abnormal Psychology.* New York: Holt, Rinehart and Winston, Inc., pp. xii+450, S5.95.

A. P. Goldstein (1966). *Psychotherapy and the Psychology of Behavior Change.* New York: John Wiley, pp. ix+472, 68s.

J. C. Gowan and G. D. Demos (1966). *The Disadvantaged and Potential Dropout.* U.S.A.: Charles C. Thomas, pp. xix+624, $18.50.

C. E. Turner (1966). *Planning for Health Education in Schools.* London : Longmans, Unesco, pp. ix+157, 17s. 6d.

E. Douvan and J. Adelson (1966). *The Adolescent Experience.* New York : J. Wiley, pp. xii+471, 60s.

L. A. Dwight (1966). *Modern Mathematics for the Elementary Teacher.* New York : Holt, Rinehart and Winston, Inc., pp. viii+598., $8.50.

P. I. Jacobs (1966). *A Guide to Evaluating Self-Instructional Programs.* New York : Holt, Rinehart and Winston, Inc., pp ix+84, 12s.

D. L. Farnsworth (1966). *Psychiatry, Education and the Young Adult.* U.S.A. : C. C. Thomas, pp. xv+268, $9.50.

H. Freund (1966). *Psychopathology and the Problems of Stuttering.* U.S A. : Charles C. Thomas, pp. xx+233, $8.75.

G. E. W. Wolstenholme and M. O'Connor (1966). *Immigration Medical and Social Aspects : Ciba Foundation Report.* London : J. and A. Churchill, Ltd., pp. xii+124, 15s.

I. C. Maclean (1966). *Child Guidance and the School.* London : Methuen and Co., pp. 117, 9s. 6d.

P. B. Johnson, *et al.* (1966). *Physical Education : A Problem-Solving Approach to Health and Fitness.* New York : Holt, Rinehart and Winston, pp. xix+486 S6.95.

Gerard de Brigode (1966). *L'architecture Scolaire.* Paris : Presses Universitaires de France, pp. viii+134, no price.

N. R. F. Maier (1966). *Principles of Human Relations.* Science Ed. New York : John Wiley, pp. ix+474, 21s.

D. Wyss (1966). *Depth Psychology : A Critical History.* London : George Allen and Unwin, pp. 568, 52s.

J. W. Atkinson (1966). *A Theory of Achievement Motivation.* New York : J. Wiley, pp. x+392, 87s.

R. S. Illingworth and C. M. Illingworth (1966). *Lessons from Childhood.* Edinburgh : E. and S. Livingstone, pp. xi+384, 37s. 6d.

W. J. H. Sprott (1966). *Sociology.* London : Hutchinson and Co., pp. 192, 10s. 6d.

G. C. Helmstadter (1966). *Principles of Psychological Measurement.* London : Methuen and Co., pp. xx+248, 36s.

Arvo Lehtovaara (1966). *The Predictiveness of the Finnish Matriculation Examination : A Differential Psychological Study.* Helsinki : Suomalainen Tiedeakatemia, pp. 288, no price.

A. R. King and J. A. Brownell (1966). *The Curriculum and the Disciplines of Knowledge.* New York : John Wiley and Sons, pp. xv+221, 45s.

Donald H. Kausler (1966). *Readings in Verbal Learning.* New York : John Wiley and Sons, pp. xii+578, 60s.

J. E. McGrath (1966). *Small Group Research : A Synthesis and Critique of the Field.* London : Holt, Rinehart and Winston, pp. ix+601, 100s.

Ruthanna Penny (1966). *Practical Care of the Mentally Retarded and Mentally Ill.* U.S.A. : Charles C. Thomas, pp. xvii+252, $9.50.

Jean Guehenno (1966). *Jean-Jacques Rousseau* (2 Vols). London : Routledge and Kegan Paul, pp. xvi+416 and xi+316, 100s. Set.

AN INTRODUCTION TO EDUCATIONAL PSYCHOLOGY

E. Stones

A book for students of education presenting an integrated view of educational psychology which embraces the most recent findings in the field as well as the major aspects of previous work.

36s ; University Paperback 18s

PLEASE, MISS, CAN I PLAY GOD?

Joan Haggerty

'A direct and enthusiastic account of dramatic work with children in a Junior School in the East End of London . . . Written lightly, directly from experience, and with the kind of honesty one would wish all teachers could bring to their work.' *The Guardian.* 21s

---

MODERN TEACHING

A new series for training colleges and teachers

Edited by P. W. Musgrave

The first title in the series:

CHILD GUIDANCE AND THE SCHOOL

Isabella C. Maclean

This book is designed to increase understanding of the problems faced in school by both normal and abnormal children. The author discusses school problems and specific types of difficult children, making valuable suggestions of ways in which school and home co-operation can be effective in the prevention of the difficulties so often encountered by psychologists, parents and teachers. 18s ; Paperback 9s 6d

---

The Library of Educational Thought

GEORG KERSCHENSTEINER

His thought and its relevance today

Diane Simons

An informed account of the modern German system of vocational education, pioneered by Georg Kerschensteiner. The success of his system and its relevance to British industrial problems and to the implementing of the Industrial Training Act, make this book essential reading for everyone concerned with education in this country. 25s ; Paperback 12s 6d

METHUEN

THE BRITISH JOURNAL OF PSYCHOLOGY

Edited by ARTHUR SUMMERFIELD

Contents of Volume 57, Parts 1–5, 1966

Satinder, K. Paul. Effects of intermodal stimulation of figural after-effects.

Kinsbourne, M., Warrington, Elizabeth K. and James, Merle. Uncertainty and transitional probability in the span of apprehension.

Pritchatt, D. and Holding, D. H. Guiding Deutsch's model in reverse.

Matthews, W. A. Sutherland's two-stage theory of discrimination learning and its experimental support.

Dale, H. C. A. Positive and negative recency in two-choice guessing.

Woodhead, Muriel M. Varying the number of alternatives in short-term recall.

Anthony, W. S. Working backward and working forward in problem-solving.

Wetherick, N. E. The inferential basis of concept attainment.

Tajfel, Henri and Bruner, Jerome S. The relation between breadth of category and decision time.

Dawes, Robyn Mason. Memory and distortion of meaningful written material.

Das, J. P. Relation between semantic satiation and verbal conditioning.

Kennedy, R. A. and Keene, A. G. The effect of short periods of food-deprivation on human performance.

Eysenck, Hans J. and Thompson, Warren. The effects of distraction on pursuit rotor learning, performance and reminiscence.

Eysenck, H. J. and Willett, R. A. The effect of drive on performance and reminiscence in a complex tracing task.

Clark, A. D. B. and Cooper, G. M. Age and perceptal-motor transfer in imbeciles: task complexity as a variable.

Clarke, Ann M., Cooper, G. M. and Henney, A. S. Width of transfer and task complexity in the conceptual learning of imbeciles.

Hasan, Parween and Butcher, H. J. Creativity and intelligence: a partial replication with Scottish children of Getzels' and Jackson's study.

Burt, Cyril. The genetic determination of differences in intelligence: a study of monozygotic twins reared together and apart.

Bell, C. R. and Watts, Anne N. Personality and judgments of temporal intervals.

Boe, Erling E. and Kogan, William S. Social desirability in individual performance on thirteen MMPI scales.

Heim, A. W. and Watts, K. P. The Brook reaction test of interests.

Mair, J. M. M. Prediction of grid scores.

Jahoda, Gustav. Geometric illustrations and environment: a study in Ghana.

Pastore, Nicholas and Terwilliger, Marlene. Induction of stereoscopic depth effects.

Locks, Erwin. The contradiction of epiphenomenalism.

Publications Reviewed.

Subscription 90s. net per volume of 4 parts

Orders may be placed with a bookseller or sent direct to the publishers

CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200 Euston Road, London, N.W.1

THOMAS BAN

Conditioning and Psychiatry

The last decade has seen a marked awakening among psychologists and psychiatrists to the significance of Pavlov's teachings, particularly in their relation to therapy. *Conditioning and Psychiatry* is the first comprehensive review of Soviet and East European as well as Western research on classical and modern Pavlovian conditioning theory and technique and their application to clinical diagnosis and therapy. 40s.

DIETER WYSS

Depth Psychology

A CRITICAL HISTORY

The critical aspect of this work is quite simply an attempt to establish the position of depth psychology today, seventy years after its inception. 52s.

JOHN COHEN

A New Introduction to Psychology

This book has been prepared as an up-to-date text especially suitable for students in Teachers' Training Colleges. It provides a general introduction to the basic problems, methods and contributions of contemporary psychology and devotes several chapters to the psychological foundations of educational practice—in the study of child development, learning and remembering, thinking and communicating. 25s. *cloth*, 12s. 6d. *paper*.

LEONARD BLANK

Psychological Evaluation in Psychotherapy

Written by a man who is both a therapist and a psychodiagnostician, this is the first book to document the direct relationship between a patient's response patterns on a battery of psychological tests and his subsequent behaviour in psychotherapy and it makes a major contribution both as a manual and a reference book for clinical psychologists, psychiatrists, social workers, trainees and students. 70s.

MARTHE ROBERT

The Psychoanalytic Revolution

SIGMUND FREUD'S LIFE AND ACHIEVEMENT

Freud's achievement lies in the fact that psychoanalysis is today an unavoidable part of modern knowledge. In *The Psychoanalytic Revolution* Marthe Robert retraces the lonely road which led Freud to his extraordinary discoveries and, through this long struggle to impose his ideas, the history of the liberating movement he founded. 50s.

PER SAUGSTAD

An Inquiry into the Foundations of Psychology

Dr. Saugstad's dominant interest has been in the area of thinking. Many psychologists will be familiar with his published work in this field. To gain a clearer understanding of the thought processes, he has carried out extensive studies of perception. This book originated in an attempt to reconcile a phenomenological and behaviouristic approach to psychology and it presents a conception of psychology which does take into account man's cognitive processes. 18s. 6d.

GEORGE ALLEN & UNWIN

40, Museum Street, London, W.C.1

SCHOOLS FOR YOUNG OFFENDERS

Gordon Rose

A comprehensive account of the development and structure of the approved school system in England and Wales, with some discussion of its methods and problems.

International Library of Criminology No 17. 52s. net

ABSENT

School Refusal as an Expression of Disturbed Family Relationships

Max B. Clyne

'This book is addressed to general practitioners, teachers and others who work with children. It gives a very full account of the problems of school refusal and it is rewarding to examine the case records of one practitioner whose experience of this often complicated disorder is extensive.' Edmund A. Harvey-Smith, *World Medicine.*

'This is a useful and readable medical-sociological study.' *The Times Educational Supplement*

38s. net

STUDIES OF TROUBLESOME CHILDREN

D. H. Stott

A challenging re-examination of the causes of maladjustment that offers consistent evidence based on extensive case material for a common congenital factor underlying behaviour disturbance.

'The empirical findings offered in this book cannot be disregarded. They constitute a challenge to much current thinking and are of importance to all who are concerned with maladjusted behaviour in any of its forms.' *The London Teacher*

32s. net

PSYCHOTHERAPY AND CHILD DEVELOPMENT

An Introduction for Students

Jean Biggar

A child psychiatrist discusses, with a wealth of clinical examples, the formative part played by the experiences of the early years in establishing the foundations of character and of mental health.

'. . . will be of great interest to doctors, teachers, clergy, social workers and students.'

British Book News

25s. net

THE DEVELOPMENT OF SEX DIFFERENCES

Edited by Eleanor E. Maccoby

This book deals with differences between the sexes in behaviour other than specifically sexual behaviour. The focus is on how sex differences in temperament and ability develop in young children, and on possible explanations for these differences.

Stanford Studies in Psychology April 60s. net

TAVISTOCK PUBLICATIONS

THE BRITISH JOURNAL OF MEDICAL PSYCHOLOGY

Edited by Dr. THOMAS FREEMAN

Volume 40, *Part* 1. CONTENTS *March*, 1967

MARTIN M. WHITTET. Highland and Island Reflections.
WAUD H. KRACKE. The Maintenance of the Ego.
DAVID M. KISSEN. Psychological Factors, Personality and Lung Cancer, in Men aged 55-64.
JOHN DENFORD. Depressive Reactions of Short Duration.
LOUIS H. SIGLER. Emotion and Athersclerotic Heart Disease.
A. D. FORREST. The Differentiation of Hysterical Personality from Hysterical Psychopathy.
JOHN E. GEDO. On Critical Periods for Corrective Experience in the Therapy of Arrested Development.
ANDREW C. SMITH. Clinical Notes on Pinel's " Treatise on Insanity."
HERBERT S. STREAN. Rationality and Irrationality in Psychotherapy : Reflections on Psycho-Analytic Technique.
Reviews.

Subscription £4 10s. per volume of 4 parts. Single parts 30s. plus postage

Orders may be placed with any bookseller, or direct to :

CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200, Euston Road, London, N.W.1

AUSTRALIAN JOURNAL OF PSYCHOLOGY

Editor : I. K. WATERHOUSE, Macquarie University, N.S.W.

Vol. 18 *No.* 3 *December*, 1966

HAMMOND, S. B. An Interactionist Approach to Personality.
O'BRIEN, G. E. Behaviourism and Explanation.
STONE, ANTHONY R., IMBER, STANLEY D , and FRANK, JEROME D. The Role of Non-Specific Factors in Short-Term Psychotherapy.
ANDERSON, C. C., and CROPLEY, A. J. Some Correlates of Originality.
RADCLIFFE, J. A. WAIS Factorial Structure and Factor Scores for Ages 18 to 54.
WALKER, WENDY L. A Preliminary Investigation of Personality Change in Depressed Patients Following ECT.
RUSSELL, P. N. and GREGSON, R. A. M. A Comparison of Intermodal and Intramodal Methods in the Multidimensional Scaling of Three-Component Taste Mixtures.
FUREDY, J. J. Reaction Time as an Index of Masking and the Effect of Check Trials on Thoughtful Subjects.
OVER, R. Context and Movement as Factors Influencing Haptic Illusions.
STANLEY, GORDON and FUREDY, J. J. Size Constancy and Emmert's Law of Apparent Sizes.
KEARNLEY, G. E. Hue Preferences as a Function of Ambient Temperatures.
IVINSKIS, A. A Note on the Open-Field Test of Emotionality
YATES, AUBREY J. A Note on Progressive Matrices (1962).
HARRISON, R. J. Predictive Validity of Some Adjustment Tests : A Research Note.
PATON, I. F. A Note on a Simple Method of Making Lip-Synchronized Sound Film.

The Journal is published three times a year. Subscription rates 36s. (sterling), and $5·50 in U.S.A. Members of B.P.S. and A.P.A. should enquire about special subscription rates from the Secretary of their Society. Non-members should send enquiries to Melbourne University Press, Parkville, N.2., Victoria, Australia, or to its agents.

in U.K.
Cambridge University Press,
200 Euston Road,
LONDON, N.W.1

in U.S.A.
Cambridge University Press,
32 East 57th Street,
NEW YORK 10022, N.Y.

Enquiries from the Orient should be sent directly to Melbourne University Press.

What Psychologists Say About the Latest MMY

THE SIXTH MENTAL MEASUREMENTS YEARBOOK

Edited by OSCAR KRISEN BUROS

▶ " most recent and definitive consumers' report of the psychometric world. This edition for the first time lists all the tests that are currently in print. * The 396 contributing authorities read like a Who's Who in the world of testing. * the *Mental Measurements Yearbooks* certainly are the test user's best defense against inadequate tests and hard selling test salesmen. Every counselor who makes use of tests and every professional person who is involved with building testing programs should have this book available to him. * If you have been using the *Mental Measurements Yearbooks* before, nothing needs to be said except, ' He has done it again.' If you are not familiar with this book, insist to your principal or supervisor that it must be available for your use."—BUFORD STEFFLRE, *Personnel and Guidance Journal*, April 1966

▶ "*The Sixth Mental Measurements Yearbook* is a 1,714 page collection of reviews of tests and books which is as indispensable to personnel psychologists as Webster's dictionary, Roget's thesaurus, and the APA directory. There can be no question about the usefulness of the volume * In 8 of the 14 reviews of the Fifth Yearbook (which—with characteristic completeness—Professor Buros has included in the Books and Review section of the current Yearbook) the adjective *monumental* is used in reference to the Yearbook. This continues to be highly appropriate."—THOMAS L. HILTON, *Personnel Psychology*, Autumn 1966

▶ "An evaluation of this last volume can be couched, with palpable redundancy, in terms applied by the reviewers of previous volumes in the series : This remarkable, monumental, colossal, and elaborate book, representing gargantuan and heroic efforts applied compulsively and with missionary zeal is certain to become a ' must '—perhaps even the indispensable ' bible '—for the increasing number of test users. * In relation to general price increases the Yearbooks are still ' best buys '."—DOROTHY C. ADKINS, *Science*, March 18, 1966

▶ "An editorial masterpiece and a beautifully printed volume which is a tribute to the assiduous efforts expended by Oscar Buros over a period of nearly six years, *The Sixth Mental Measurements Yearbook* is, as one would expect from the Buros tradition, the best one to date."—JOAN J. BJELKE and WILLIAM B. MICHAEL, *Educational and Psychological Measurements*, Summer 1966

▶ " For over a quarter of a century psychologists have relied heavily upon the *Mental Measurements Yearbooks* for information about tests they wish to use. This new volume will immediately and inevitably find its place alongside the earlier ones. These do not, of course, become obsolete, but are supplemented by the latest volume and with it and the general index, *Tests in Print* (1961) by the same editor, form a unique and comprehensive body of knowledge about all known tests published in English. * tremendous service to psychometry."—L. B. BIRCH, *British Journal of Educational Psychology*, June 1966

Published November, 1965, 1752 2-*column pages*, $32.50.

THE GRYPHON PRESS
220 Montgomery Street, Highland Park, New Jersey 08904, U.S.A.

Studies in Child Development

This series of books, published by Longmans, in association with the National Bureau for Co-operation in Child Care, has a twofold object : to make the results of research available to a wide public as rapidly as possible ; and to disseminate, with a minimum of technical jargon, research findings concerned with the development and needs of children.

FOUR YEARS ON

Stan Gooch & M. L. Kellmer Pringle

A follow-up study at school leaving age of children formerly attending a traditional and a progressive junior school. *17s. 6d. net*

January 1967

ADOPTION–FACTS & FALLACIES

M. L. Kellmer Pringle *with the assistance of* **Micheline Dewdney, Eileen Crellin** *and* **Rosemary Dinnage**

A review of research in the United States, Canada and Great Britain, between 1948 and 1965.

11,000 SEVEN-YEAR-OLDS

M. L. Kellmer Pringle, N. R. Butler and R. Davie

The first report of the National Bureau Child Development Study (1958 Cohort) submitted to the Central Advisory Council for Education (England).

DEPRIVATION AND EDUCATION

M. L. Kellmer Pringle

'*This book should be prescribed reading for all who work professionally with young children whether deprived, maladjusted or normal. It is a treasure trove of processed thought.*' Stephanie M. Leese. BRITISH JOURNAL OF PSYCHIATRY. *42s. 6d.*

1724

Longmans

The British Journal of
Social and Clinical Psychology
Edited by MICHAEL ARGYLE and R. R. HETHERINGTON

Contents of Vol. 5, Part 4, December, 1966

TAYLOR, R. L. Differences in accuracy of public and private judgments.
WARR, P. B. and KNAPPER, C. The role of expectancy and communication content in indirect person perception.
WARREN, N. Social class and construct systems : an examinaton of the cognitive structure of two social class groups.
SHERWOOD, J. J. Authoritarianism, moral realism, and President Kennedy's death.
LYNN R. Brainwashing techniques in leadership and child rearing.
RAO, S. A note on the investigation of the MPI with different occupational groups in India.
PARKE, R. D. and WALTERS, R. H Alcoholism, avoidance learning and emotional responsiveness.
HUNDLEBY, J. D. The construct validity of a scale of acquiescence.
FRANKS, C. M., and MANTELL, D. Introversion-extraversion and the verbal conditioning and generalization of meaning responses to nonsense syllables in normal and alcoholic subjects.
FARLEY, F. H. Individual differences in solution time in error-free problem solving.
MUKHERJEE, B. N. 'Social desirability' and 'anxiety' variables in three measures of anxiety.

The subscription price per volume of four parts is 90s. net. Orders may be sent to your bookseller or direct to the publishers

CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200, Euston Road, London, N.W.1

OCCUPATIONAL PSYCHOLOGY
Editor : ALEC RODGER

OCTOBER 1966 VOLUME 40, No. 4

CONTENTS

Psychology and Military Proficiency. By N. A. B. WILSON.
Concepts of Work. By SYLVIA SHIMMIN.
Some Characteristics of Foremen in the Iron and Steel Industry. By PETER B. WARR, MICHAEL BIRD and RAYMOND W. HADFIELD.
Frequent Alteration of Shifts on Continuous Work. By J. WALKER.
Age and Internal Labour Mobility of Semi-Skilled Workers. By C. K. ELLIOTT.
Age, Speed and Accuracy : A Study in Industrial Re-Training. By G. H. JAMIESON.
Book Reviews.

Annual Subscription 40 *shillings, overseas* 50 *shillings*

National Institute of Industrial Psychology
14, Welbeck Street, London, W.1.

OTHER PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology.
The British Journal of Psychology (Monograph Supplements).
The British Journal of Medical Psychology.
The British Journal of Mathematical and Statistical Psychology.
The British Journal of Social and Clinical Psychology.

The British Journal of Psychology and *The British Journal of Medical Psychology* appear in quarterly parts in volumes of about 350 pages, subscription price of each, payable in advance, 60s. per volume (post free). Subscriptions may be sent to any Bookseller or to the Cambridge University Press, Bentley House, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $10.00. Enquiries should be addressed to the Cambridge University Press, American Branch, 32, East 57th Street, New York 22.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November, in volumes of about 184 pages. The subscription rate payable in advance is 30s. 6d. per volume (post free). The subscription price in the U.S.A. is $4.50 (post free). Subscriptions should be sent to the Society's Agents, Taylor and Francis, Ltd., Red Lion Court, Fleet Street, E.C.4, or to any Bookseller.

The British Journal of Social and Clinical Psychology is issued three times a year in February, June and November, in volumes of approximately 240 pages. The subscription price of each, payable in advance, is 50s. (post free). Subscriptions may be sent to any Bookseller, or to the Cambridge University Press, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $8.50. Enquiries should be sent to the Cambridge University Press, American Branch, 32, East 57th Street, New York 22. Separate parts will be sold at 20s. : $3.50 in the United States.

A series of *Monograph Supplements* is issued in connection with the *British Journal of Psychology*, but is not included in the subscription for that journal. Orders may be sent to any Bookseller or to the Cambridge University Press, 200, Euston Road, London, N.W.1, from whom details of these publications may be obtained.

MANUSCRIPTS

For the *British Journal of Psychology* should be sent to Professor ARTHUR SUMMERFIELD, Department of Psychology, Birkbeck College, Malet Street, London, W.C.1.

For the *British Journal of Medical Psychology*, should be sent to Dr. THOMAS FREEMAN, Lansdowne Clinic, 4, Royal Crescent, Kelvingrove, Glasgow, C.3.

For the *British Journal of Educational Psychology*, should be sent to Mr. L. B. BIRCH, Institute of Education, The University, Sheffield, 10.

For the *British Journal of Mathematical and Statistical Psychology*, should be sent to Dr. R. J. AUDLEY, Department of Psychology, University College, Gower Street, London, W.C.1.

For the *British Journal of Social and Clinical Psychology*, should be sent to Mr. MICHAEL ARGYLE (Social Psychology Editor), Institute of Experimental Psychology, 1, South Parks Road, Oxford;

or to:

Dr. RALPH HETHERINGTON (Department of Studies in Psychological Medicine), 6, Abercromby Square, Liverpool, 7.

For the *Monograph Supplements*, to Professor IAN M. L. HUNTER, Department of Psychology, University of Keele, Keele, Staffs, whose decision regarding the publication of all *Monograph Supplements* will be final. No MS. can be considered for publication in any year unless it is in the hands of the Editor by 1st September of the preceding year.

OFF-PRINTS

Contributors receive 50 free copies of their article and may buy more in units of 50.

Members of the British Psychological Society receive all the Society's Journals at reduced rates.

Information on membership and further activities may be obtained from the Secretary, Tavistock House South, Tavistock Square, London, W.C.1.

PUBLISHED BY THE ASSOCIATION OF TEACHERS IN COLLEGES AND DEPARTMENTS OF EDUCATION

Education for Teaching is published three times a year at mid-term in November, February and May, and distributed free to members of the Association. To non-members it costs 5s. per issue, post free (subscription 14s. per annum).

Members of the A.T.C.D.E. receive the *British Journal of Educational Psychology* at a reduced rate.

Information on membership and further activities may be obtained from the Honorary Secretary, 151, Gower Street, London, W.C.1.

CONTENTS

1 NOV 1968

H. Lytton : Follow-up of an Experiment in Selection for Remedial Education . 1

D. F. Swift : Family Environment and 11+ Success : Some Basic Predictors . 10

D. S. Finlayson and L. Cohen : The Teacher's Rôle : A Comparative Study of the Conceptions of College of Education Students and Head Teachers 22

S. Mitchell and M. Shepherd : The Child Who Dislikes Going to School 32

E. D. Cave : The Effects of the 'Eleven-Plus' Result on the Subsequent Careers of Pupils 41

M. L. Peters : The Influence of Reading Methods on Spelling 47

M. D. Shipman : Environmental Influences on Response to Questionnaires . . 54

E. E. Rump and V. Southgate : Variables Affecting Aesthetic Appreciation in Relation to Age . . 58

M. P. M. Richards and H. E. Ross : Developmental Changes in Children's Drawings 73

H. J. Eysenck : Intelligence Assessment : A Theoretical and Experimental Approach . 81

K. W. Rogers : Concepts of Time in Secondary School Children of Above Average I.Q. 99

Research Notes :

W. Yule, L. Lockyer and A. Noone : The Reliability and Validity of the Goodenough-Harris Drawing Test . 110

P. Williams : Regression Analysis Programmes for the Identification of Important Variables in Educational and Psychological Experiments . 112

R. J. Goldman and D. F. Clarke : The Minnesota Tests of Creative Thinking—A Note on Scorer Reliability in Follow-up Studies with English Primary School Children . . 115

A. Abi Rafi : The Progressive Matrices (1938) and the Dominoes (D48) Tests : A Cross-Cultural Study 117

R. R. Dale : Co-education as a Possible Influence on the Desire for Marriage. . 119

J. K. Backhouse : The Use of Valentine's Reasoning Tests in an Investigation into Transfer from Mathematics to Reasoning 121

Comment :

H. Lytton : Note on the Article " Differences in Test Scores and in the Gaining of Selective Places for Scottish Children and those in England and Wales," by Douglas, Ross, Maxwell and Walker 124

Thesis Abstracts :

T. D. Swales : The Attainments in Reading and Spelling of Children Who Learned to Read Through the Initial Teaching Alphabet 126

G. Regan : Personality Characteristics and Attitude to School 127

J. A. Chadwick : Some Effects of Increasing the Teachers' Knowledge of their Pupils' Self Pictures 129

W. P. L. Johnson : The Religious Attitude of Secondary Modern County School Pupils . 132

M. A. Mycock : A Comparison of Vertical Grouping and Horizontal Grouping in the Infant School 133

M. E. Ainsworth : The Relationship Between Motivation, Personality, Intelligence and School Attainment in a Secondary Modern School 135

D. Sharples : Factors Affecting the Composition Performance of Ten-year-old Children 137

E. Soloman : Personality Factors and Attitudes of Mature Training College Students 140

Critical Notices :

C. Burt : Gifted Children, by M. Branch and A. Cash 143

R. R. Dale : The Public Schools : A Factual Survey, by Graham Kalton . . 149

Book Reviews 152

Publications Received 162

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

JUNE 1967

VOL. XXXVII, PART 2

Price One Pound

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by
L. B. BIRCH

Associate Editors:
P. E. VERNON and DORIS M. LEE

Advised by

SIR CYRIL BURT
R. A. C. OLIVER
E. A. PEEL
SIR FRED J. SCHONELL
S. WISEMAN

Vol. XXXVII, Part 2

The British Journal of Educational Psychology is issued by the British Psychological Society and the Association of Teachers in Colleges and Departments of Education, once a term, in February, June and November. The subscription is £2 a year post free, payable in advance, or £1 each number.

Subscriptions and Orders should be sent to "The Manager, B.J.E.P. Department, Methuen and Co., Ltd., 11, New Fetter Lane, London, E.C.4. Members of the British Psychological Society and members of the A.T.C.D.E. receive the *Journal* on special terms through the Societies.

Papers for publication should be sent to L. B. Birch, The Institute of Education, The University, Sheffield, 10. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for the *British Journal of Educational Psychology* should be sent to L. B. Birch, at the above address. *Advertisements* and *Subscriptions* for this *Journal* should be sent to Manager, B.J.E.P. Department, Sweet and Maxwell, Spon Booksellers, Ltd , 11, New Fetter Lane, London, E.C.4.

Published by
METHUEN AND CO., LTD., 11, NEW FETTER LANE, LONDON, E.C.4,
for the British Journal of Educational Psychology, Ltd.

Registered Offices: 42, Hill Street, Birmingham, 5. This address should *not* be used for letters or books intended for the Editor, Authors, or the Publishers.

On behalf of
The British Psychological Society and the Association of Teachers in Colleges and Departments of Education.

All rights reserved.

PRINTED BY THE BIRMINGHAM PRINTERS, LTD., 42, HILL STREET, BIRMINGHAM, 5.

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

Edited by
L. B. BIRCH

Associate Editors:
P. E. VERNON and DORIS M. LEE

Advised by
SIR CYRIL BURT
R. A. C. OLIVER
E. A. PEEL
SIR FRED J. SCHONELL
S. WISEMAN

Vol. XXXVI
1966

Published by
METHUEN AND CO., LTD., 36, ESSEX STREET, STRAND, LONDON, W.C.2.,
for the British Journal of Educational Psychology, Ltd.
Registered Offices: 42, Hill Street, Birmingham, 5. This address should *not* be used for letters or books intended for the Editor, Authors, or the Publishers.

On behalf of
The British Psychological Society and the Association of Teachers in Colleges and Departments of Education.

All rights reserved.

PRINTED BY THE BIRMINGHAM PRINTERS, LTD., 42, HILL STREET, BIRMINGHAM, 5.

VOLUME XXXVII JUNE 1967 PART 2

THE ENGLISH PROFICIENCY OF OVERSEAS STUDENTS

By ALAN DAVIES
(*Department of Applied Linguistics, University of Edinburgh*)

SUMMARY. Seven tests of English language reception including listening comprehension, reading comprehension, reading speed and grammar, were tried on samples of overseas students in the U.K. and overseas, and on a control group of native English speaking students. Three factors extracted by principal components were tentatively identified as segmental listening, textual listening (over + ·6) and a more diffuse reading comprehension (over + ·5). Concurrent and predictive validity studies gave coefficients which ranged from + ·8 to −·2. Analysis of variance indicated that home language area both by itself and in interaction with sex has considerable influence on English proficiency. Comparative figures suggested that students requiring most remedial attention are: Semitic, Mediterranean, West African and East/Central Africa. It was found that varying academic levels (e.g., university, technical college) did not require varying proficiencies in English and that one proficiency cut off was adequate.

I.—INTRODUCTION.

ABOUT 50,000 (or one-tenth) of all students in Britain in higher education come from overseas countries where English is not the first language. They study in all kinds of institutions all over Britain; but most (about 34,500) are in the London area. Most of them are not here primarily to study the English language but something else through the medium of English. It seems likely that the proportion will continue (Robbins, 1963, App. 1, p. 165).

The research described in this paper was set up jointly by the British Council and Birmingham University to investigate the difficulties met by overseas students using English as a second language and, if possible, to assess the level of English proficiency required by overseas students. The ultimate purpose was to provide a prognostic test which could be used before final decisions were taken about admitting or encouraging overseas students to pursue higher education in Britain.

II.—PREVIOUS RESEARCH.

Published research on language test construction is scanty. Lado's work applied one branch of American Linguistics to the testing of language proficiency and, as a result, set up a theory of foreign language testing. This theory may be summarised: a learner has language problems; find the problems: test them. The theory assumes that testing control of the problems is testing control of the language (Lado, 1961, p. 24). The problems are found by the methods of *contrastive analysis*, of comparing the learner's first language with his target language, system by system, thereby exposing the learner's difficulties, i.e., the places where the two languages do not coincide. Contrastive analysis as a necessary preparation for language testing has been challenged on theoretical grounds (Upshur, 1962; Carroll, 1961, George, 1962); and in any case, the real need in English proficiency testing is for a test which can be applied to multi-language backgrounds.

There is other and more recent American work, by no means all in Lado's image. What is lacking in the published reports (Harris, 1961, 1964) is information about predictive validity; what is given is based on existing tests of English proficiency.

As far as the *components* of English language proficiency are concerned, Agard and Dunkel's (1948) investigation distinguished between the reading and the oral/aural skills. In a later study (with. however, particular reference to first language English speakers), Spearitt (1962) found that listening comprehension formed a separate factor in language control. Several writers have claimed that proficiency varies according to academic level and subject (Palmer, Slocum, Mackenzie, 1961 ; Carroll, 1961 ; Agard and Dunkel, 1948). Contrastive analysis has played a major part in the literature (Coppock, 1961 ; Lado, 1961 ; Strevens, 1961) but other variables such as sex, age, health and so on have been posited as bearing on proficiency (Lado, 1961 ; E.T.I.C., 1962).

III.—Test construction and sampling.

This research started with two assumptions, one linguistic, the other psychometric. The linguistic starting point was, with Lado, that it was indeed language problems that should be tested but that these should be located in areas of performance rather than in points of contrast between languages. The psychometric starting point was that a language proficiency test should have high predictive power. What matters to an overseas student is whether he is academically successful in his own subject, and whether he has enough English for this end ; not whether he has a near native control over English.

In this enquiry it was hoped to find out : first, whether it is possible to construct an English language proficiency test which has satisfactory correlations with measures of academic success ; second whether different academic levels (e.g., university post-graduate, technical colleges) require varying proficiencies in English, ; third, whether such variables as age, sex and home language area have any influence on English proficiency.

Tests of language production are notoriously difficult to construct and to use; it was decided, therefore, to concentrate on reading comprehension and listening comprehension and to draw on two areas for the content of the battery. The first area was the linguistic, based on linguistic categories ; the second was language at work, based on work samples of what overseas students have to use their English for. The eventual make up of the final trial battery was as set out in Table 1.

TABLE 1

Composition of Test Battery.

Test	Items	Content	Skill	Aspect
1	65	Phonemic Discrimination : words in isolation	Listening.	Linguistic.
2	25	Phonemic Discrimination : words in sentences	Listening.	Linguistic.
3	50	Intonation	Listening.	Linguistic.
4a	8	Listening Comprehension : general	Listening.	Work sample.
4b	5	Listening Comprehension : specialised science	Listening.	Work sample.
4c	5	Listening Comprehension : specialised : arts	Listening.	Work sample.
5	196	Reading Speed	Reading.	Work sample.
6a	49	Reading Comprehension : general	Reading.	Work sample.
6b	50	Reading Comprehension : specialised arts	Reading.	Work sample.
6c	50	Reading Comprehension specialised science	Reading.	Work sample.
7	50	Grammar	Reading.	Linguistic.

For the final trial three separate groups were postulated:

(1) Overseas students already studying in Britain, if possible in their first year (UKOS), N=496.

(2) Native English speaking students (NE) N=267.

(3) Overseas students in their own countries (OSOS) N=238.

An attempt was made to draw the UKOS samples from representative language groups and types of institution. The NE group consisted of college of education students and post-graduate diploma of education students. The OSOS group was drawn where local British Council officials and university authorities were willing and able to help and were mainly university students in nine countries in Asia, Africa and the Middle East.

Two sets of criteria were used to assess the validity of the trial-battery: first, students' grades or examination results; and second, teachers' or tutors' estimates of their students' English proficiency, based both on a rank order and on placement on a defined five-point scale. The first of these was intended as a predictive validity criterion, the second as a concurrent validity criterion.

IV.—RESULTS.

Item Analysis. Item Analysis (extreme groups method) was performed in the UKOS sample (N=120) and a total of twenty-two items rejected (out of the total Battery of 500 items). The critical validity index was taken as +·20. The faulty items all came from three tests, Nos. 1 (Phonemic Discrimination); 3 (Intonation) and 7 (Grammar).

Reliability. Reliability for the Battery is set out in Table 2 along with Means and Standard Deviations.

Factor Analysis. The results of a Factor Analysis (principal components, manually rotated) of the nine tests done by all students are presented in Table 3.

Validity. Correlation coefficients were obtained for thirty-one sets of students. These are presented (along with a list of the students' institutions and centres attended) in two sets in Tables 4 and 5, which show the coefficients (scaled standard scores for each test summed, but not weighted) and for Short Version (Tests 1, 3, 6a, 7), also scaled to standard scores, summed and not weighted.

V.—DISCUSSION.

Item Analysis. The Test (Test 1) of Phonemic Discrimination had a number of faulty items · almost certainly this was a *noise* problem. The most faulty a test (Test 3) was, as expected, that of Intonation and Stress in Conversation: there seem to be three main reasons, all linguistic: native speakers' intonation operates in different systems; ambiguity is not fully realised phonologically; a test of intonation alone omits the essential redundancy of e.g., paralinguistic features.

TABLE 2

RELIABILITY, MEANS AND STANDARD DEVIATIONS.

	Reliability	Items	Mean			SD			SE of Measurement
	Split half corrected N=250	No.* of Items	UKOS N=496	NE N=267	OSOS N=238	UKOS	NE	OSOS	UKOS
Test 1	·91	58	41·8	52·4	36·5	10·1	3·1	8·2	3·03
2	·85	25	17·3	23·6	14·6	4·9	1·9	5·1	1·96
3	·75	38	21·2	30·9	19·1	5·3	4·9	5·2	2·65
4a	·51	8	4·1	5·8	4·4	1·6	1·3	1·4	1·12
4b		5	2·7	4·0	3·0	1·2	0·3	0·3	·84
4c		5	1·6	2·8	1·6	1·2	1·3	1·1	·84
5	·97	196	70·1	136·4	83·3	33·6	38·9	34·3	6·72
6a	·94	49	27·5	39·6	32·4	13·2	8·5	11·8	3·40
6b	·88	50	18·4	21·2	20·5	7·9	6·2	7·2	2·76
6c	·93	50	26·2	35·8	26·9	10·6	8·6	9·2	2·75
7	·89	47	33·6	44·8	36·7	8·5	2·4	7·3	2·80
Total Scaled Score	—	—	80	102·6	79·7	12·6	8·0	10·6	—
Short Version Scaled Score Total	—	—	40	50·1	39·5	6·0	3·4	5·2	—

* After Item Analysis.

TABLE 3

FACTOR ANALYSIS (PRINCIPAL COMPONENTS).

Tests	Factors (Rotated)		
	A	B	C
1	·781	·300	—·328
2	·800	·221	·049
3	·695	·145	·223
4a	·007	·877	·163
4b	·223	·617	·250
5	·577	·066	·570
6a	·524	·279	·580
7	·580	·237	·513
4c	·348	·291	·600

TABLE 4

VALIDITY : CONCURRENT VALIDITY.

5 pt. Scale and R.O.

	Centre	Rank Order of Correlation Coefficients		
		N	Total Score	4 Tests
1	Birmingham University : Commonwealth Teacher Bursars	24	·813	·837
2	Birmingham · British Council mixed group	10	·704	·669
3	Birmingham University : mixed group	30	·685	·594
4	Selly Oak Colleges · adult education	24	·645	·464
5	Leicester University Commonwealth Teacher Bursars	17	·596	·603
6	London : British Council mixed group	55	·550	·558
7	Thurrock Technical College . mixed group	23	·403*	·196*
8	Bromsgrove College of Further Education · mixed group	12	·360*	·170*
9	Leeds University : mixed group	32	·230*	·175*
10	Manchester University : mixed group	27	·083*	·075*
	Combined Totals (by Fisher's z)	254	·525	·475
	Overseas			
1	N. Nigeria : University mixed group	23	·802	·809
2	Iraq : British Council mixed	8	·631*	·535*
3	Thailand : medical group	24	·624	·545
4	Malaya . Teacher Training College	22	·529	·396
5	Uganda : British Council mixed	20	·246*	—·070*
6	Malaya : University mixed group	16	—·240*	—·212*
	Totals (by Fisher's z)	113	·510	·420
	Overall Totals	367	·520	·455

* Not significant at 5 per cent. level.

Reliability. Except in the case of Test 4 (three subtests of Listening Comprehension), the split half (corrected) coefficients were considered satisfactory. The low Listening Comprehension coefficient is in line with earlier findings (Spearritt, 1962 ; Agard and Dunkel, 1948). What is noticeable here is that Test 3, another test of Listening Comprehension (but of a more segmental variety), does achieve satisfactory reliability. This seems to agree with Spearritt (1962, p. 66) " tests involving a considerable amount of auditory synthesis may tend to be less reliable than those in which the speech sample is clearly presented."

Factor Analysis. The Factor Analysis suggested the existence of three factors, tentatively identified as Segmental Listening, Textual Listening and a more general Reading Comprehension. Test 4a, a Listening Test, was found in the third factor, that of General Reading Comprehension. It was considered that its highly literary content and its somewhat involved questions caused its association in the General Reading Comprehension Factor. The solution (Table 3) is very tentative and mainly useful for the immediacy of the distinction

TABLE 5

VALIDITY : PREDICTIVE VALIDITY.

Grade/Exam.

Centre	Rank Order of Correlation Coefficients		
	N	Total Score	4 Tests
1 Hall Green Technical College . Special English Course	9	·743	·731
2 Bristol University : Commonwealth Teacher Bursars.	22	·678	·642
3 Birmingham Coll. Commerce : Special English Course	30	·672	·586
4 London British Council : mixed group	10	·643	·612*
5 Bolton Technical College : mixed group	31	·624	·637
6 Leeds University : mixed group	22	·477	·556
7 Manchester University : mixed group	12	·459*	·633
8 Birmingham University : mixed group	5	·381*	·501*
9 Brooklyn Technical College : mixed group	22	·292*	·306*
10 Bromsgrove Coll. Further Education : mixed group	12	·126*	·087*
11 Birmingham College of Commerce : commercial course	18	—·019*	—·192*
12 Thurrock Technical College : mixed group	15	—·197*	—·219
Totals (by Fisher's z)	208	·460	·455
Overseas			
1 Uganda : British Council mixed	12	·669	·513*
2 Malaya · Teacher Training College	22	·470	·515
3 N. Nigeria · University nixed group	18	·364*	·415*
Totals (by Fisher's z)	52	·485	·480
Overall Totals	260	·465	·460

* Not significant at 5 per cent. level.

it makes between the two listening factors. It is between these rather than between Listening and Reading that whatever polarity there is in the Battery exists.

Validity. The range of validity coefficients (Tables 4 and 5) is large ; from +·8 to —·2, but in all cases where the coefficient is below +·4, the N is too small to make the correlation significant. It was considered, therefore, that this English Language Proficiency Test Battery showed concurrent and predictive validity. Satisfactory concurrent validity was established in a variety of situations (both in the UKOS sample) and overseas (in the OSOS sample) for centres containing university students, both graduate and postgraduate, Commonwealth Teacher Bursars, adult education students and widely mixed students representing the whole range for whom the British Council is responsible A similar claim is made for predictive validity.

Influence of Variables. Simple analyses of variance were made (Tables 6, 7, 8). The analysis on Test 1 (Phonemic Discrimination) suggested that home language by itself accounts for the variation in score on this test. This was expected since it is particularly in a phonemic teaching or testing inventory that contrastive language analysis has its major contribution to make. Test 7 (Grammar) shows a similar home language influence, but this time with a

TABLE 6

ANALYSIS OF VARIANCE.

Test I. UKOS.

Language Area and Sex.

Source of Variation	df	Sums of Squares	Mean Square	F
Between groups—language areas. . .	10	329·75	32·975	4·828*
Sex	1	0·24	0·24	·003
Interaction	10	41·18	4·118	0·602
Error	474	—	6·83	—
Total	495	371·17	—	—

* Significant at the 1 per cent. level.

TABLE 7

ANALYSIS OF VARIANCE.

Test 7 : UKOS.

Language Area and Sex.

Source of Variation	df	Sums of Squares	Mean Square	F
Between groups—language areas.	10	154·99	14·59	2·88*
Sex	1	·007	·007	·001
Interaction	10	103·003	10·30	1·92*
Error	474	—	5·37	—
Total	495	258	—	—

* Significant at the 1 per cent. level

TABLE 8

ANALYSIS OF VARIANCE.

Total Score : UKOS.

Language Area and Sex.

Source of Variation	df	Sums of Squares	Mean Square	F
Between groups—language areas.	10	42852·9	4285·29	375·5*
Sex	1	3·8	3·8	0·33
Interaction	10	18032·7	1803·27	158·0*
Error	474	—	11·41	—
Total . . .	495	60889·4	—	—

* Significant at the 1 per cent. level.

possibly significant sex interaction. When the different aspects of proficiency tested are summarised, in Total Score (Table 8), the language sex interaction becomes highly significant, with, of course, the home language still showing a significant F ratio.

The findings of these simple analyses were, therefore, that neither sex nor age, either alone or in interaction, had any effect, but that home language area, in interaction with sex, had a considerable influence on test scores.

Of the eleven language areas delineated, the two West European groups (Romance and Germanic), the Creolised and the two Eastern groups (India and Pakistan ; and N. and S.E. Asia) all have means above the general mean (Table 9).

TABLE 9

STATISTICS FOR UKOS : HOME.
Language Area Subgroups.

Subgroup	N	Mean	S.D.
1. Germanic	66	87·7	9·79
2. Creole (W. Indies and Sierra Leone)	39	84·6	14·73
3. North and South-East Asia ..	44	83·1	9·77
4. India and Pakistan..........	73	80·9	13·42
5. Romance	40	80·2	14·56
6. West Africa	56	77·6	10·82
7. Slav	22	77·5	12·08
8. Mediterranean	31	77·4	8·52
9. East and Central Africa......	46	76·4	11·40
10. Semitic	58	73·9	13·27
11. Francophonic	21	70·5	6·81
Total	496	80·0	12·6

From a simple investigation of the means alone, then, it is the two African groups (West and East/Central), the Mediterranean group, the Semitic and the Slav groups who show poorest performance and, hence, greatest deficiency. And of these five groups, the Middle East Semitic group is poorest, superior only to the third language Francophonic speakers. The other four are very close together, with the East/Central African group weakest, but only by very little. It is, perhaps, the Mediterranean group of Turks and Greeks who cause most concern since they have so small an ability range. Excluding the Slav group, whose numbers are few, it is students from these four areas who, on the basis of this evidence, need most attention, the Semitic, Mediterranean, West African and East/Central African.

The Indian group has a very central mean, but its wide ability-range suggests that some of the students from India and Pakistan need as much attention as those in the four groups mentioned above. This would be in line with the Michigan conclusions on their Indian students at Ann Arbor (University of Michigan, 1959).

Decision Making. Using expectancy tables, it was estimated that on Total Score (Mean=80; SD=12·6) a cut-off point of 72 would provide a nearly 80 per cent. success for the measure. For the Short Verson (Mean=40 : SD=6) a score of 36 would provide a similar (slightly higher) percentage success. Differing cut-offs for various levels and types were tried; but all the evidence pointed to one single-cut-off for all overseas students. (Omitted from this result were those students actually studying the English language itself, since with their great variety of thresholds a similar variety of entry points would be needed.)

Once again there was a difference between concurrent and predictive validity. Both for Total Score and for Short Version, the optimum predictive cut-off was lower than the concurrent (68 : 72 ; 34 : 36). This suggested that teachers demanded a higher English proficiency than was needed for examination purposes. It may be that this teacher expectation is unrealistic and somewhat irrelevant. Further indication of this was given when the concurrent criteria (i.e., the teachers' estimates, etc.) were themselves tried as predictors (Davies, 1965, p. 185). They were no more satisfactory than the tests. But, however important the end criterion, such as an examination, the fact remains that teachers have to teach their overseas students and need to be assured that their proficiency is adequate. It, therefore, seems sensible to accept the concurrent cut-off score as the one single optimum cut-off for all overseas students.

VI.—IMPLICATIONS.

There are three chief implications of this enquiry. First, *validity theory*. A valid English proficiency test can be constructed having satisfactory relationship to the desired criterion (end of year examination or grade in student's own subject of study). More imporant, the enquiry suggested that *content* validity is not an adequate measure of a language test's worth. There must be an attempt to grapple with the terminal behaviour which it is desired to predict and somehow sample this for test purposes This may be done (as here) through *predictive* validity where the terminal behaviour is performance in a task (i.e., an examination) which partly involves performance in English. It could even better be done through *construct* validity which would take for its construct the terminal behaviour postulated.

Second, *cut-off* " Setting a cutting score requires a value judgment " (Cronbach, 1960, p.334). Whatever the optimum cut-off score found in this battery it must be the decision of a receiving institution where to draw its own line. This decision must be based on a realistic description of the language demands made on its own students by their academic courses. The tester needs the descriptive linguist to tell him just what is the universe of discourse to be sampled.

Third, *proficiency standards and demands*. The finding that one proficiency standard is adequate for success at all academic levels needs further well-designed experimentation. If the observational findings of this enquiry (viz., that *one* standard is adequate) were substantiated, this could have an important influence on the teaching of English to foreign students. At the same time, both the decision making analysis and the validity study of the Short Version suggested that the demands made by teachers and tutors on their overseas students may be too high. It could be that it is native adequacy rather than efficient performance that is being set up as the model.

VII.—Acknowledgements.

The writer wishes to express his thanks to Proffessor E. A. Peel for super vision during this research and to Mr. G. E. Perren, then of the British Council, among others, for advice and help.

VIII—References.

Agard and Dunkel (1948). *An Investigation of Second Language Teaching.* Boston : Ginn.

Carroll, J. B. (1961). Some fundamental considerations in testing for English language proficiency of foreign students. In *Testing.* CAL, Washington D.C. : pp. 30-40.

Coppock, H. (1961). *Testing Skills in Spoken and Written English.* Working Paper for Commonwealth Conference, Makerere (unpublished mimeo).

Cronbach, L. J. (1961). *Essentials of Psychological Testing.* New York : Harper and Row, 2nd Ed.

Davies, A. (1965). *Proficiency in English as a Second Language.* Unpublished Ph.D. thesis, University of Birmingham.

Etic (1962). *English of Overseas Students in Britain.* English Teaching Information Centre, British Council, London, mimeo.

George, H. V. (1962). Testing—another point of view. *English Language Teaching,* Vol. 16, No. 2, pp. 72-78.

Harris, D. P. (1961). The American University Language Centre Testing Program. In *Testing.* CAL Washington D.C., pp. 41-58.

Harris, D. P. (1964). Report on Feb. 17th, 1964 Administration of TOEFL, unpublished mimeo. Washington D.C. : Center of Applied Linguistics.

Lado, R. (1961). *Language Testing.* Longmans.

Mackenzie, N. (1961). English proficiency testing in the British Commonwealth. In *Testing.* CAL, Washington D.C. : pp. 54-72.

Palmer, L. A. (1961). The Michigan Overseas Testing Service. In *Testing.* CAL, Washington D.C. : pp. 73-84.

Robbins, Lord (1963). *Report of Committee on Higher Education,* Appendix 1. London : H.M.S.O.

Slocum, J. B. (1961). Assessing the English proficiency of foreign students. In *Testing.* CAL, Washington D.C. : pp. 97-101.

Spearritt, D (1962). Listening comprehension—a factorial analysis. *ACER Research Series No.* 76. Melbourne.

Strevens, P. D. (1961). Objective testing (mimeo). Also in *Papers in Language and Language Teaching.* OUP 1965, pp. 87-103.

Upshur, J. A. (1962). Language proficiency testing and the contrastive analysis dilemma. *Language Learning.* Michigan : Vol. 12, No. 2, pp. 123-127.

University of Michigan (1959). *A Comparison of Scores on the E.L.I. English Proficiency Test with First-Semester Grade Point Averages for 599 Foreign Students at the University of Michigan.* E.L I., University of Michigan, Ann Arbor, 1959.

(*Manuscript received 27th January,* 1966)

DISCRIMINATION LEARNING AND DISCRIMINATION LEARNING SETS IN SUBNORMAL CHILDREN*

By E. A. LUNZER and I. HULME
(*Department of Education, University of Manchester*)

Summary. The aim of the enquiry was to establish whether subnormal children have more difficulty than normals of comparable mental age in learning a visual discrimination and in forming a discrimination set. Twenty-eight older sub-normal children and twenty-four normal pre-school children were required to solve six visual discrimination problems. Owing to imperfect matching, the normal group had a somewhat higher mean M.A. Nevertheless, little difference appeared either in the number of trials to master the first problem or in subsequent improvement attributable to the learning set.

It was found that children who were rated as lacking in self-confidence fared badly in the experiment.

Spontaneous naming of the discriminanda was as frequent in the subnormal group as it was among the normals. Indeed, questioning children at the end of a problem as to their mode of solution elicited significantly more such ' discriminating verbalisation " from the subnormals than it did from the normals. Nevertheless, such language appeared to be unrelated to the speed of learning.

A comparison of the present results with others reported in the literature suggests that specific deficit is more likely to be found at higher mental ages when the role of language may be more decisive.

I.—Introduction.

Since the publication of Harlow's (1949) article on the formation of learning sets, a number of writers (Hunt, 1961 ; Mandler, 1962) have begun to look for the possibility of a coherent account of the development of intellectual processes in which the formation of sets gives rise to the emergence of new strategies, and these in turn allow the operation of new ' structures.' Inevitably, such a view must be regarded as problematic. However, the research worker who inclines to it is apt to pay particular attention to processes involved in the formation of learning sets. So far as the learning of the subnormal is concerned, a new interest will be attached to the question whether subnormal children are less capable of forming learning sets than normal children who have reached an equivalent level of development, even though they may be equally capable of learning a specific task. If this proved to be correct, we could readily account for their failure to maintain the same rate of progress as normal children.

Two questions need to be answered. First, are subnormal children as well able to learn visual discriminations as normal children ; and, second, are they inferior in the acquisition of discrimination learning sets?

On both questions, the literature is divided. Taking first the question of visual discrimination learning, subnormal children were found to be poor in studies by House and Zeaman (1958a, 1958b), Zeaman, House and Orlando (1958), and Stevenson and Iscoe (1955). On the other hand, studies by Stevenson and Zigler (1957), and O'Connor and Hermelin (1963, pp. 63-4) showed no significant differences between sub-normals and controls. Similarly, while there have been fewer studies of the formation of discrimination learning sets in the subnormal, their results are again at variance with one another. Both Kaufman and Peterson (1958) and Girardeau (1959) found significant differences

---

* This research was facilitated by a grant to the senior author from the National Society for Mentally Handicapped Children.

in favour of normals, while an earlier study by Plenderleith (1956) found none. Ellis (1958) and Stevenson and Swartz (1958) found higher grade subnormals to be superior to low-grades.

Plenderleith, and also Kaufman and Peterson, used a multiple-problem procedure to investigate the formation of learning sets. The former allowed six trials per problem., while the latter allowed only three. There is still some uncertainty as to the effect of reducing the number of trials per problem to the point where transition from problem to problem occurs before the first problem has been solved to criterion. Although Harlow, Meyer and Settlage (1953), maintain that the ideal number of trials per problem should be low (probably less than twenty-four, and possibly less than six), at least for the formation of an oddity learning set, Hayes, Thompson and Hayes (1953) set out from the contrary view that " Learning set is a function of the number of problems learnt, in which case a problem would be wasted if it were dropped before it was learned, and trials would be wasted if continued beyond mastery " (p. 99). Since several investigations, including those cited above, have shown the formation of learning sets under conditions which do not permit the solution of the earlier problems in the series, the view of Hayes, Thompson and Hayes is probably too extreme. Nevertheless, if subnormal children are, in fact, poorer than children of comparable mental age in this area, considerable importance attaches to the question whether the weakness lies in the discrimination learning itself, or in the formation of the discrimination LS.

Girardeau's study is instructive. Trials were continued on each of five problems to a criterion of eleven consecutive correct choices within a single fifty-trial session. Girardeau found differences in favour of his controls on both counts. Girardeau used children having a mean mental age which was somewhat over 5 (and his control subjects had a slightly higher mean mental age than his group of mongoloids). Also the discrimination problems used in this study were difficult. Both these factors may have contributed to his results. There is good reason to believe that a qualitative change in this type of learning occurs about the age of 5–6, when even the first problem is solved within a very few trials, and one-trial learning seems to be the rule for all subsequent problems, with little fluctuation (Hayes, Thompson and Hayes, 1953). Such a change may be associated with the verbal direction of behaviour (Luria, 1961; Kendler and Kendler, 1962). Similarly, it has been said that perceptual processes are poorly developed in mental defectives, unless trained (Clarke and Blakemore, 1961).

It was, therefore, decided to compare the formation of discrimination learning sets in sub-normal children below the mental age of 5, with that of normal children of similar mental age, using discriminanda which, although two-dimensional, were perceptually very different. An attempt was made to record all spontaneous language behaviour.

II.—PROCEDURE.

The apparatus consisted of six pairs of cardboard boxes 2½ inches square and 1 inch high. Only the lids of these boxes were used and these were decorated with prominent colour-form diagrams, consttutinig the discriminanda. These were presented to the child on a tray. For each pair of boxes, constituting the stimulus elements for one problem, the diagrams differed both in colour and in shape. Problem 1 contrasted a green hollow square with a red circle. The remaining problems in the order 2–6 were: yellow star *versus* black triangle, pink X *versus* large blue square, green ring *versus* red N, pink shape in the form of two smaller partially superposed squares *versus* blue diamond, red horizontal line *versus* blue cross. For each problem, the reward was hidden under the same stimulus for all trials, while the position of the stimuli was varied in accordance with a predetermined random sequence. The child was required to pick up one of the boxes only and find the sweet. A non-correction procedure was used throughout: if the child failed to make a correct choice, a screen was interposed and the next trial was prepared by re-arranging the boxes. Sweets were used as rewards. Since the subject could be expected to acquire many sweets even without solving the problem, being rewarded on average 50 per cent. of the time, the sweets were collected by the child and dropped into a narrow test tube, so that he could observe the rate at which his reward was accumulating. It was found that motivation could be maintained at a high level, using this procedure (up to thirty trials per day, this figure having been adopted as a limit to avoid losing the child's interest).

For each problem, trials were continued until the subject reached a criterion, which was set at eight out of ten successive correct choices, subject to the following condition: that the number of errors (0, 1 or 2) must be less than the number of occasions upon which the child continued to choose the correct object even though the position of the boxes had been changed. This criterion is somewhat more generous than that used by many others in this field, but was adopted after a great deal of preliminary experimentation, in the course of which it was found that, whether from inattention, or from a desire to vary the situation, children often made a wrong choice, even after it was obvious to the experimenter that they had solved the problem.

A limit of sixty trials (two sessions) was allowed for the solution of each problem excluding the criterion trials. Children who failed to solve the first problem within this number of trials were not given the remaining problems. However, experimentation was continued with children who solved the first problem even if they failed to solve one or more of the later problems. In order to obtain the maximum transfer of learning set, the first problem was re-administered on the session following its solution (using either the same criterion, or the shorter criterion of five correct choices) immediately before presenting the second problem. Similarly, the second problem was given again before presenting the third, and the third before the fourth. The fourth, fifth and sixth problems were given only once.

To meet the administrative requirements of the people in charge of the centres in which the experiment took place sessions were held at intervals of three–four days for both normal and subnormal subjects. Although several subjects were seen on each day, in no case was a subject given two sessions on the same day. For each subject, once experimentation had begun, sessions were continued until the entire programme was completed, so that inter-session intervals never exceeded three–four days. For all sessions, except the first, the subject was first given five trials of the problem which he had solved in the previous session, and this was followed immediately by the next problem on the list. When the previous problem had not been solved, experimentation in this problem was continued, and the next problem was introduced only in the session following. In any one session, only one fresh problem was learnt.

III.—SUBJECTS.

The subnormal group consisted of twenty-eight children. Of these, twelve subjects, whose ages range from 5 : 1 to 6 : 6, were attending an observation school for young subnormal children in Manchester. The measured I.Qs. of these subjects ranged from 46–62. The remaining sixteen subjects were taken from a junior day training centre, also in Manchester. Their ages ranged from 6 : 2—14 : 9 with measured I.Qs. between 35 and 58. Ten of the twenty-eight subjects were mongols; in many cases the aetiology of the remainder was unknown, and no single category was predominant.

The control group consisted of twenty-four children ranging from 2 : 7—4 : 4 in chronological age. These were drawn from three of Manchester's nursery schools. Members of this group were not pre-selected for I.Q. All of the children aged 3+ were taken, together with three who were just over 4 and one who was under 3, but unusually co-operative. Since I.Qs. were not available for these children at the time of selection, intelligence tests were carried out some months later to allow an estimate of their M.A. at the time of the experiment. It will be apparent from Table 1 that the mean M.A. of the control group turned out to be significantly higher than that of the subnormals, the difference representing approximately eight months for the two complete samples and just under seven months for the two groups of eighteen subjects who completed the learning series. (It will be seen that six normal subjects and ten subnormal subjects failed to complete the first problem within sixty trials and were dropped from the remainder of the enquiry).

TABLE 1

SUBJECTS: C.A., I.Q. and M.A.*

(*a*) The Subnormal Sample.
(Figures in second row are children who completed the main series.)

N	C.A.		I.Q.		M.A.	
	Mean	Range	Mean	Range	Mean	Range
28	8 : 2	5 : 1—14 : 7	47	30—62	3 : 8	2 : 6—5 : 11
18	8 . 8	5 : 1—14 : 7	48	35—62	3 . 7½	2 . 8—5 . 11

(*b*) The Normal Sample.

N	C.A.		I.Q.		M.A.	
	Mean	Range	Mean	Range	Mean	Range
28	3 : 7	2 : 7—4 : 2	116	94—139	4 : 4	3 : 2—5 : 4
18	3 : 7	2 : 7—4 : 2	118	99—139	4 : 5	3 : 2—5 : 11

(*c*) Difference in M.A. (years).

Normals		Sub-normals		Difference		
Mean MA	N	Mean M.A.	N			
4·33	24	3·64	28	0·68	3·30	(P< ·01)
4·39	18	3·83	18	0·56	2·05	(P< ·05)

---

* We are indebted to Miss O. C. Sampson who arranged for much of the testing to be carried out by her students, being post-graduates in psychology preparing for clinical work.

IV.—RESULTS.

(1) *The acquisition of discrimination and discrimination learning set.*

Six of the twenty-four normal subjects failed to solve the first discrimination problem within the sixty trials allotted. Of the twenty-eight subnormal subjects, ten failed the first problem. The number of subjects who participated in problems 2–6 was, therefore, eighteen in both groups.

Table 2 affords a summary of the number of trials taken to criterion on successive problems by the subjects in the two groups.

TABLE 2

NORMAL AND SUB-NORMAL GROUPS TRIALS TO CRITERION FOR SUCCESSIVE PROBLEMS.

		(*a*) Normal Subjects.						(*b*) Sub-normal Subjects.					
	Problems	1	2	3	4	5	6	1	2	3	4	5	6
Trials to Criterion	0—1	2	7	6	10	9	12	3	5	5	9	9	14
	2—5	2	5	3	1	4	2	1	7	3	2	3	1
	6—20	9	2	5	6	3	2	8	2	5	5	3	1
	21—60	5	—	3	—	1	—	6	3	3	—	1	—
	Failed*	6	+4	+1	+1	+2	+1	10	+1	+2	+2	+2	+2

* Children who failed to reach criterion on the first problem were not set the remaining problems.

It is characteristic of this type of performance that the number of subjects reaching criterion within brackets of trials representing any constant interval decreases progressively from zero e.g. more problems are solved in 0–2 trials than in 3–5, more in 3–5 than in 6–8, etc. In order to minimise the resultant skewing of scores (trials to criterion), these were first converted to log scores (to base e, using the Manchester Atlas computer), and the calculations of Tables 3 and 4 are based on the converted scores. Correct first guesses (0-trial learning), were treated as instances of 1-trial learning, with a converted score of 0.

There is a decline in score for successive problems presumably due to the formation of the learning set, and this is apparent in both groups. Differences between the mean scores attained by the normal and sub-normal groups do not attain significance for any of the problems. It is apparent also that differences in problem difficulty, which might have been eliminated in the experimental design, contribute some unevenness to the results. Problems 2 and 4 appear to have proved easier than the rest. However, this effect was identical for both groups.

Learning effects were tested separately for the two groups, following the method of analysis of variance outlined by MacNemar (1957). This allows one to test for linearity of regression of score on problem order.

TABLE 3

NUMBER OF TRIALS TO CRITERION FOR SUCCESSIVE PROBLEMS.
(18 normal and 18 sub-normal subjects.)

		Order of Problem					
		1	2	3	4	5	6
Mean . . .	Normal . . .	15·11	16·06	12·22	7·95	7·01	7·95
	Sub-normal	18·06	10·95	16·89	9·89	11·50	7·45
Mean of converted Scores	Normal . . .	2·37	1·51	1·66	1·09	1·02	0·77
	Sub-normal	2·41	1·50	2·04	1·17	1 28	0·63
s.d. (converted Xcores)	Normal	1·02	1·63	1·39	1·37	1·27	1·34
	Sub-normal	1·22	1·33	1·43	1·42	1·48	1·32
t	—	0·09	0·01	0·80	0·15	0·55	0·30
Difference*	—	0·03	0·00	0·38	0·07	0·26	—0·13

* Positive values signify differences in favour of the normal group, i.e., fewer trials to criterion.

TABLE 4

ANALYSIS OF VARIANCE.

(*a*) Normal Subjects.

Source	df	Mean Square	F
Linear regression	1	25·962	13·77**
Between problems	5	5·983	3·11**
Within problem arrays	102	1·922	—
Deviation from linearity	4	1·412	—
Residual from linear regression	106	1·886	<1

Value of r : —·339 (df 106).
Value of η : —.364 (df 102).

(*b*) Sub-normal Subjects.

Source	df	Mean Square	F
Linear regression	1	27·975	14·04**
Between problems	5	7·325	3·69**
Within problem arrays	102	1·986	—
Deviation from linearity	4	2·930	1·48 (NS)
Residual from linear regression	106	1·992	—

Value of r : —·342 (df 196).
Value of η : ·391 (df 102).

(Linear regression is tested against residual from linear regression between problems, representing the correlation ratio, and deviation from linearity, against within problem arrays.)

The results for the two groups are similar. Linear regression (and coefficient of correlation) is significant at the ·01 level. We may, therefore, conclude that there is evidence of learning set. Other inter-problem differences already referred to are not significant, and we cannot reject the hypothesis of linearity: i.e., since the scores have been converted to a logarithmic scale, that improvement is a logarithmic function of order of problem.

However, although significant, the correlation is not high. This reflects the continuance of considerable fluctuation in performance over succeeding problems as is apparent in the following examples.

TABLE 5

TRIALS TO CRITERION IN SUCCESSIVE PROBLEMS.

	Subject	MA .	Problem					
			1	2	3	4	5	6
6 normals	HN	3 : 2	7	0	1	1	0	0
	HB	3 : 9	0	62	8	2	12	3
	SP	3 . 10	18	5	3	1	0	1
	SV	4 : 0	16	0	17	11	2	0
	JD	4 1	5	5	2	1	0	0
	RT	5 : 11	18	0	0	0	0	0
6 Sub-normals	MS	3 : 3	15	0	21	0	1	1
	CE	3 · 4	5	3	0	1	0	0
	CI	3 · 6	26	2	0	5	27	0
	AL	3 . 6	19	4	0	6	6	0
	JP	3 : 7	16	5	23	6	13	0
	JS	4 : 5	9	0	27	0	5	6

(2) *Relation between visual discrimination learning and confidence.*

In the course of the experiment, it was noted that children who were less self-assured in the test situation, seemed to have considerable difficulty in solving the first problem. It was, therefore, decided to test whether this impression could be sustained by a slightly more systematic enquiry.

The teacher in charge of each of the five centres (or schools) concerned was asked to rate each of the children under her care on a seven-point scale ranging from ' extremely confident ' to ' extreme lack of confidence.' Teachers were encouraged to use the whole of the scale. We would, of course, stress that the ratings were made without knowledge of the children's performance in the learning situation and on the basis of general behaviour in class.

The original impression is borne out by the figures shown in Table 6 ' Failures ' include children who failed on all subsequent problems after solving the first as well as those who failed to solve the first problem. All of the subjects who were rated in the lowest category failed in the experiment. Of the six normal failures, only one appears as average in the teacher's rating. Seven of the twelve subnormal failures were rated below average in confidence, three as average and only two above. Table 7 combines the data for the two groups, with the ratings pooled to form three categories. The relation between confidence and success proves highly significant: $\chi^2 = 13{\cdot}83$ (2 df, $P < {\cdot}001$).

TABLE 6

CONFIDENCE RATINGS OF CHILDREN FAILING OR SUCCEEDING IN VISUAL DISCRIMINATION LEARNING.

Normal Subjects.

	Rating						
Success or Failure	1	2	3	4	5	6	7
S	3	0	4	8	0	2	0
F	0	0	0	1	2	1	3

Subnormal Subjects.

	Rating						
Success or Failure	1	2	3	4	5	6	7
S	0	3	5	4	2	2	0
F	0	0	2	3	0	5	2

TABLE 7

GROUPED CONFIDENCE RATINGS FOR CHILDREN FAILING OR SUCCEEDING IN VISUAL DISCRIMINATION LEARNING.

	Rating		
Success or Failure	1, 2 or 3	4	5, 6 or 7
S	15	12	6
F	2	4	13

$\chi^2 = 13 \cdot 83$.

(3) *Language.*

A number of studies (e.g., Cantor, 1955; Dietze, 1955; Weir and Stevenson, 1959) have shown the value of language in the acquisition of visual discrimination. In the main, these findings relate to a somewhat different type of situation from the present, in that the discriminanda were, in fact, difficult to discriminate so that the learning consisted either wholly or partly in gains in the ability to attend to those cues which differentiated the stimulus objects. However, the possibility remains that spontaneous naming accelerates visual discrimination learning even when the discriminanda differ markedly from one another, as in the present investigation. Certainly, the naming of objects, by shape or by colour, should favour an object-choice strategy (which is the correct strategy in terms of the experiment) as against a positional strategy. In particular, it was thought possible at the outset that the normal subjects would perform better than the subnormals and, if such proved to be the case, there was reason to believe that this higher level of performance might be a function of superior language ability It is well-known that subnormals tend to be specifically retarded in certain linguistic functions (Thompson and Magaret, 1947 ; Luria, 1961 ; O'Connor and Hermelin, 1963 ; Hulme, 1965).

No attempt was made in the present investigation to discover the effect of a specific instruction to name the two objects before making a choice. However, so far as possible, a complete record was taken of all the language used by the subjects in the course of the experiment (even though the record is far from complete for these children who chattered a lot), and care was taken to include in these records every instance in which the subject named or otherwise referred to the choice-subjects In addition, every time a subject reached criterion on a problem, he was asked how he knew which one to choose. If he merely pointed or gave no reply, or if he made some irrelevant remark, he was pressed a little further (even though in many cases such questioning did not lead to a better answer). The replies to these questions were recorded separately.

From the point of view of the role of spontaneous verbalisation in this type of learning, the crucial language responses are those in which the subject names one or both objects or describes them in some way, e.g., " This is a star," " It's (the reward is) under the red," etc. Such utterances have been termed discriminating remarks.

There is clearly no significant difference in the performance of the two groups (Table 8).

TABLE 8

INCIDENCE OF DISCRIMINATING REMARKS.

	Normal (N=24)	Sub-normal (N=28)
Total incidence of DR	39	44
Number of subjects offering at least one spontaneous	15	17
Incidence of D.Rs. in session where subject failed to reach criterion . ..	12	5
Number of subjects offering spontaneous D.R. and failing to reach criterion	4	4

In the vast majority of cases, the subject eventually reached criterion in the course of the session in which the remark was made. But this does not imply that the discrimination helped him to reach criterion, for in no case did a verbal discrimination herald a sudden insight.

Table 9 is a breakdown of the replies given to the observer's questions upon successful completion of each of the six problems by the eighteen normals and eighteen subnormals. For the first five columns, the value of χ^2 is 20·7 (3 df), significant at the ·001 level. A 2×2 comparison of the first column and the

TABLE 9

REPLIES TO QUESTIONS PUT BY OBSERVER.

	D.R.	V.P.	M.V.	P	O	Total—D.R.
Normal Subjects	11	32	40	7	14	93
Subnormal Subjects	33	14	27	8	15	64

D.R. = discriminating replies
V P. = pointing with verbal accompaniment (e.g. "It was under that one ")
M.V = miscellaneous verbal replies (e.g. " I found it")
P. = points
O. = no reply

last (i.e., discriminating replies *versus* all others), yields a χ^2 of 16·1 (1 df), which again is significant at the ·001 level. Eight normal subjects offered a discriminating reply on at least one occasion, as compared with fourteen of the subnormals. Again the difference favours the subnormals though χ^2, at 2·9 (Yates' correction) is not significant ($\cdot 05 < p < \cdot 10$). These results have to be weighed against the previous finding that no important differences appeared in discrimination and there was some advantage to the normals.

V.—DISCUSSION.

(1) The principal aim of the enquiry was to determine whether significant differences exist between normal and sub-normal children in visual discrimination learning, in the formation of learning sets, or in both.

Although the findings are not entirely unambiguous, they do permit of a fairly definite conclusion : within the limits of the situation described in this enquiry, there are some advantages to normal children, but these are slight and usually insignificant. Neither the failure to learn visual discriminations nor the failure to form elementary learning sets can account for their intellectual retardation at this level of development.

No less than ten sub-normal subjects fail to master the first problem within sixty trials. This result is comparable with House and Zeaman's (1958a) finding that many sub-normals failed in a visual discrimination learning task, although all of twelve monkeys had succeeded in the investigations reported by Harlow (1944, 1949, 1951). But it is more favourable ; for in House and Zeaman's study only seven out of thirty-seven were successful, whereas in our study ten out of twenty-eight failed. There were differences in the number of trials allowed (fifty in the House and Zeaman study), and there may well have been differences in motivation. The crucial differences is that no controls were made in their study, and the authors are content to argue from the analogy of a necessarily different comparative study. In the present study normal children were given exactly the same tests. Six out of twenty-four normal children also failed to master the original discrimination. The difference is clearly not significant, and moreover the control group had a slight advantage in mental age.

Of the remaining eighteen subjects in both groups, very few failed to solve any of the succeeding problems. In both groups a large majority of subjects solved the second problem in far fewer trials than the first. Moreover, there was continued improvement in both groups. In general, the improvement was comparable.

Finally, in both groups fluctuation continued at least until the sixth problem in the series. The presence of such fluctuation in the performance of children in this type of task had already been noted by Hayes, Thompson and Hayes (1953).

Thus, the chief conclusion of this study is negative. The slow learning of the severely sub-normal is not typified in the visual discrimination learning situation, and must, therefore, be sought elsewhere.

However, there are two restrictions to the generality of this conclusion and both are important. The first relates to the character of the situation, and the second to the age of the subjects.

In the present enquiry, the choice objects were relatively easy to discriminate. There is reason to believe that difficult discriminations may yield different results. Thus, Girardeau (1959) found more considerable differences in favour of the normals, using more difficult discriminations.

However, by far the more important restriction to the present conclusion concerns the mental age of the subjects. Thus, even in the study by Hayes, Thompson and Hayes (1953), there are indications that beyond a mental age of 6, children form a learning set rapidly, and maintain a level very close to 1-trial learning without fluctuation for subsequent problems. This indication is fully borne out in the case of normal 11-year-olds who formed one of the groups in the study by Stevenson and Swartz (1958).

It is possible that even sub-normal subjects, or some of these, show a similar qualitative change about the mental age of 6, so accounting for the intermediate results of the higher grade sub-normals studied by these last authors, and also for the superiority of higher grade sub-normals found by Ellis (1957). On the other hand, both these results suggest that beyond the ages studied in the present enquiry, improvement in the sub-normal is more gradual and performance more erratic.

(2) It was found that the spontaneous naming of the discriminanda has little relation to the regularity of correct choice. At the same time, such naming was actually more frequent in the sub-normal than in the normal. One may not infer that the language of the sub-normal is better than or comparable with that of normal children at similar mental ages. The evidence reviewed elsewhere (Hulme, 1965), and especially the recent work of Bryant (1964) is clearly to the contrary. What results is rather that the mere naming of object represents a different level of language behaviour from that which underlies ease of learning, at a later stage of development. A similar distinction is made by Kuenne (1946) in conjunction with her study of transposition. At the later level, behaviour is 'dominated' by verbal interpretation, (c.f., Luria's conception of ' regulation '), while at the present level language is no more than an alternative response mode which subsists alongside overt behaviour. What this means is that, when compared with children of similar mental age, sub-normal children are not necessarily deficient in language of the first sort, but they are deficient in the integration of language and behaviour.

(3) Finally, a word must be said about what is learned in this type of discrimination learning. The fact that children who have not solved a given problem will quite often name the discriminanda shows that it is not the visual discrimination that is being learned.

Zeaman, House and Orlando (1958) also reject this interpretation on experimental grounds. Similarly, they reject the interpretation that what is being learned are ' the rules of the game,' on the grounds that children who fail with more abstract (geometrical) discriminations may ' play the game ' correctly with meaningful objects, yet they fail to transfer this behaviour. We would interpret this as meaning that at this level, ' playing the game ' consists in the unconscious operation of a preferred schema · the subject is not aware of the schema, and he can neither report on it nor transpose it at will.

The conclusion of these authors is that what is learned consists in ' distinctive, cue-producing responses.' Insofar as this implies that in the process of learning, behaviour is cued to specific aspects of the stimuli, we would agree. But the interpretation goes further in that it implies that the perception of the stimuli is a separate process which precedes the mediating, response-producing cue behaviour. Such an account is closely modelled on Hullian theory. An alternative way of looking at it is that the act of perception itself is always a scanning for cues, i.e., it always occurs in the context of some behavioural strategy. The behaviour to which it gives rise is not necessarily mediated by an abstraction of the cue, as in a verbal ' cue response ' occurring after the perception. Nor does the solution of the problem require such abstraction at this level. It entails no more than the substitution of one strategy for another, both of which are sensorimotor, equally available to individuals of various species,and available at an early age. Restle (1958, 1962) has offered a number of models which would account for such substitution.

VI.—REFERENCES.

BRYANT, P. (1964). The effect of verbal instruction on transfer in normal and severely subnormal children. *J. Ment. Defic Res.*, 8, 35-43.

CANOT, G. N. (1955). Effects of three types of pre-training on discrimination learning in pre-school children. *J. Exp Psychol.*, 49, 339-342.

CLARKE, A D. B., and BLAKEMORE, C. B. (1961). Age and perceptual motor transfer in imbeciles. *Brit. J. Psychol.*, 52, 125-131.

DIETZE, D. (1955). The facilitating effect of words on discrimination and generalisation. *J. Exp. Psychol.*, 50, 255-260.

ELLIS, N. R. (1957). Object quality discrimination learning set in mental defectives. *J. Comp Psychol.*, 50, 79-81.

GIRARDEAU, F. L. (1959). The formation of discrimination learning sets in mongoloid and normal children. *J. Comp. Psychol* , 52, 566-570.

HARLOW, H. F. (1944). Studies in discrimination learning by monkeys. I. *J. Gen Psychol* , 30, 1-13.

HARLOW, H. F. (1949) The formation of learning sets. *Psychol Rev.*, 56, 51-65.

HARLOW, H. F. (1951). Primate learning. In STONE, C P. (Ed), *Comparative Psychology*. New York : Prentice Hall.

HARLOW, H. F., MEYER, D., and SETTLAGE, P. H (1951) The effect of large cortical lesions on the solution of oddity problems by monkeys. *J Comp. Psychol.*, 44, 320-326.

HARLOW, H. F., and WARREN, J. M. (1952). Formation and transfer of discrimination learning sets *J. Comp. Psychol.*, 45, 482-9.

HAYES, K. J., THOMPSON, R., and HAYES, C. (1953). Discrimination learning sets in chimpanzees. *J. Comp. Psychol.*, 46, 99-104.

HOUSE, B. J., and ZEAMAN, D. (1958a). Visual discrimination learning in imbeciles. *Amer. J Ment. Defic.*, 63, 447-452.

HOUSE, B. J., and ZEAMAN, D (1958b) A comparison of discrimination learning in normal and mentally defective children. *Child Developm.*, 29, 411-6.

HULME, I. (1965). A comparison of the play, language and reasoning of severely subnormal children and children of similar mental age. *Unpubl. M.Ed. Thesis*, Univ. of Manchester.

HUNT, J. McV. (1961). *Intelligence and Experience*. New York : Ronald Press.

KAUFMAN, M. E., and PETERSON, W. M. (1958). Acquisition of a learning set by normal and mentally retarded children. *J. Compi Psychol.*, 51, 619-621.

KUENNE, M. R. (1946). Experimental investigation of the relation of language to transposition behaviour in young children. *J Exp. Psychol.*, 36, 471-490.
LURIA, A. R. (1961). *Speech and the Regulation of Behaviour*. London : Pergamon Press.
MCNEMAR, Q. (1957). *Psychological Statistics* (2nd Edit.). London . Wiley.
MANDLER, G. (1962). From association to structure. *Psychol. Rev.*, 69, 415-427.
O'CONNOR, N., and HERMELIN, B. (1963) *Speech and Thought in Severe Subnormality*. London : Pergamon Press.
PLENDERLEITH, M. (1956). Discrimination learning and discrimination reversal learning in normal and feeble-minded children. *J. Genet Psychol.*, 88, 107-112.
RESTLE, F. (1958) Towards a quantitative description of learning set data. *Psychol. Rev.*, 65, 77-91.
RESTLE, F (1962). The selection of strategies in cue learning. *Psychol. Rev* , 69, 329-343.
STEVENSON, H. W. and ISCOE, I. (1955). Transposition in the feeble-minded *J.Exp. Psychol* , 49, 11-15.
STEVENSON, H. W., and SWARTZ, J. D. (1958). Learning set in children as a function of intellectual level. *J. Comp. Psychol.*, 51, 755-7.
STEVENSON, H W., and ZIEGLER, E. F. (1957). Discrimination learning and rigidity in normals and feeble-minded individuals. *J. Pers.*, 25, 699-711.
THOMPSON, C. W., and MAGARET, A. (1947). Differential test responses of normals and defectives. *J. Abnorm. Soc. Psychol.*, 42, 284-293.
WEIR, M. W., and STEVENSON, H. W. (1959). The effect of verbalisation in children's learning as a function of chronological age. *Child Developm.*, 30, 143-9.
ZEAMAN, D., HOUSE, B. J., and ORLANDO, R. (1958). Use of special training conditions in visual discrimination learning with imbeciles. *Amer. J. Ment. Defic* , 63, 453-9.

(*Manuscript received 5th April*, 1966)

INTELLIGENCE AND FAMILY SIZE, 1949-1965

By J. D. NISBET and N. J. ENTWISTLE
(Department of Education, University of Aberdeen)

SUMMARY. An age-group of 2,868 Aberdeen school children has been tested at ages 7, 9, 11 and 12, on verbal reasoning, non-verbal reasoning and attainment tests. The relations between family size and test score were expressed as regression coefficients The usual inverse relationship between family size and test score was observed. Regression coefficients from the verbal tests were larger than from the non-verbal. Consistently larger regression coefficients were obtained for girls than for boys. Within social classes I and II, regression coefficients were smaller than within the other social classes. The results are interpreted as evidence of environmental influence on test scores. Results obtained from previous investigations in Aberdeen are compared with those obtained in the present study.

I.—INTRODUCTION.

IN *The Home and the School*, Douglas (1964), has discussed the inverse relationship between family size and intelligence. He found that this relationship persisted even when social class was held constant and suggested that this residue must be attributed to environmental influences such as deficiencies of parental care. From an earlier study, Nisbet (1953) suggested that " part (but not all) of the negative correlation of family size and intelligence test score . . . may be attributed to the environmental influence of the size of family on verbal ability and through it on general mental development." Among the results which led to this conclusion were the following :

(1) When different tests were used with the same sample, the size of the negative correlation between family size and intelligence was found to be associated with the degree to which the test measured verbal ability.

(2) As the children grew older, larger negative correlations were obtained, suggesting a cumulative environmental effect up to age 12.

The present investigation has allowed these conclusions, which were based on tests given in 1949 and 1950, to be reassessed in more detail, using an age-group tested on various occasions between 1960 and 1965.

II.—SUBJECTS, TESTS AND OTHER DATA.

The present study is based on data collected as part of an investigation into the age of transfer to secondary education in Scotland, sponsored by the Scottish Council for Research in Education. An age-group of Aberdeen children which transferred to secondary school in 1965 has been given a number of intellectual and non-intellectual tests since 1960. Data on family size and social class were obtained from the records of the Aberdeen Medical Sociology Research Unit which is conducting a parallel investigation using the same age-group. The samples for which records could be traced consisted of 2,868 pupils (1,476 boys and 1,392 girls) with test scores at ages 11 and 12, of whom 2,734 pupils (1,420 boys and 1.314 girls) had also been tested at ages 7 and 9.

The tests used were as follows : at age 7, Moray House Picture Test ; at 9, Schonell Essential Intelligence Test ; at 11, Moray House Verbal Reasoning Test 68 and N.F.E.R. Non-Verbal Test 1 ; at 12, Moray House Verbal Reasoning Test 73, Moray House English Test 35 or 36, and Moray House Arithmetic Test 35 or 36 (older children sat Tests 35). Scores on all these tests were re-standardised to means of 100 and standard deviations of 15 to allow direct comparisons between tests.

III.—ANALYSIS OF RESULTS.

Mean scores on the above tests for each family size (sexes separately) are shown below. Table 1A shows the results for the complete sample tested at ages 11 and 12. Those who also had been tested at ages 7 and 9 had mean scores on the Verbal Reasoning test at 11, which were virtually identical to those of the complete sample. Thus, the mean scores of this group at ages 7 and 9, shown in Table 1B, can be compared with those shown in Table 1A.

TABLE 1A

MEAN TEST SCORES BY FAMILY SIZE (SEXES SEPARATELY) AT AGES 11 AND 12

	Test	Verbal		Non-Verbal		Verbal		English		Arithmetic		N	
	Age	11		11		12		12		12			
	Sex	B	G	B	G	B	G	B	G	B	G	B	G
Family Size	1	104·4	105·9	102·4	103·5	103·5	105·5	104·8	107·5	104·5	103·9	145	116
	2	102·6	104·8	103·3	104·5	102·4	105·2	102·1	105·1	102·8	104·7	462	441
	3	100·7	102·6	101·8	102·8	100·3	102·7	100·3	102·2	101·3	101·9	385	379
	4	98·7	98·0	99·7	98·7	98·8	98·1	98·4	98·1	98·9	97·3	246	229
	5	94·5	97·9	95·4	98·7	93·6	97·6	93·6	97·2	95·0	97·1	134	123
	6+	93·6	91·9	94·8	92·3	93·5	92·8	93·3	92·4	95·2	92·8	104	104

TABLE 1B

MEAN TEST SCORES BY FAMILY SIZE (SEXES SEPARATELY) AT AGES 7 AND 9.

	Test	Picture		Verbal		N	
	Age	7		9			
	Sex	B	G	B	G	B	G
Family size .. .	1	104·5	107·3	104·9	107·5	138	109
	2	103·4	104·7	102·3	104·6	446	411
	3	101·9	101·7	100·8	102·3	368	359
	4	98·2	96·3	98·7	97·3	239	214
	5	94·7	96·7	94·1	96·1	129	119
	6+	91·8	92·3	92·5	91·5	100	102

The relationship between mean score on the various tests and family size can best be shown by calculating the regression coefficients. These represent the slopes of lines drawn through points plotting the change in mean test score with increasing family size. The larger the negative regression coefficient, the sharper the average drop in mean test score from one family size to another. Since the standard deviations have been made equivalent for all tests and also between sexes, the regression coefficients of test score on family size shown in Tables 2A and 2B, are directly comparable.

TABLE 2A

REGRESSION COEFFICIENTS OF TEST SCORE ON FAMILY SIZE AT AGES 11 AND 12.

Test	Verbal	Non-Verbal	Verbal	English	Arithmetic
Age	11	11	12	12	12
Boys	—2·284	—1·857	—2·067	—2·488	—2·175
Girls	—2·889	—2·376	—3·048	—3·123	—2·734
All	—2·571	—2·106	—2·514	—2·781	—2·438

TABLE 2B

REGRESSION COEFFICIENTS OF TEST SCORE ON FAMILY SIZE AT AGES 7 AND 9.

Test	Picture	Verbal
Age	7	9
Boys	—2·618	—2·219
Girls	—2·935	—2·995
All	—2·790	—2·576

The results shown in Table 2A confirm the 1953 finding that the tests involving a higher proportion of verbal items show a more marked inverse relationship with family size. There appears to be a gradient of increasing regression coefficients from the non-verbal test, through Arithmetic and verbal reasoning, to the English test. This again supports the hypothesis that part of the relationship between family size and intelligence test score is attributable to a verbal component in the test score.

A comparison between Table 2B and Table 2A shows no significant differences between the similar verbal reasoning tests taken at ages 9, 11 and 12. The regression coefficient at age 7 is actually larger than those at later ages and is much higher than the value reported in 1953. The absence of any difference in the results obtained from similar tests at various ages is in conflict with the earlier finding, which indicated an age gradient with the negative correlation between intelligence test score and family size increasing as the children became older. This was previously taken as evidence to support the hypothesis of an environmental influence extending throughout the primary school years. The present results suggest that the environmental effect must take place before the age of 7 and a similar conclusion has been reached by Douglas (1964). It is possible that the relatively standard school environment of these pupils between ages 7 and 12 prevents any increase in the intellectual differentiation between children from small families and those from large families. A parallel study by Horobin, Bytheway and Oldman (1966) is in preparation, analysing results from this same age-group by social class. The usual social class gradients in mean test scores have emerged, favouring the children from social classes I and II on the Registrar General's scale, but these differences do not increase between the ages of 7 and 12.

By analysing the data within social classes, it can be shown that the relation between intelligence test score and family size is not just another aspect of the differences between the social classes. Certainly, social classes I and II tend to produce more intelligent children and also to have smaller families, but the results shown in Tables 3A and 3B show the familiar relationship between test score and family size in all social classes for both English and non-verbal tests. As regression coefficients become unstable with smaller numbers in the sub-groups, Table 4 shows the values of the regression coefficients of English and non-verbal test scores on family size by social class with the sexes grouped

TABLE 3A

MEAN SCORES ON ENGLISH TEST AT AGE 12 BY FAMILY SIZE AND BY SOCIAL CLASS (SEXES SEPARATELY).

	Social Class (Sex)	I and II		III (Non-manual)		III (Manual)		IV and V	
		B	G	B	G	B	G	B	G
Family size.......	1	111·1	110·8	106·6	108·4*	102·9	107·4	102·2	104·8
	2	110·1	110·8	103·9	109·5	100·9	103·3	94·7	98·4
	3	106·9	107·1	102·5	104·9	98·9	102·4	96·4	97·1
	4	105·7	112·0	96·4	104·3	98·1	96·7	96·8	94·2
	5	105·9*	107·5*	89·0*	105·4*	95·6	95·4	90·1	95·9
	6+	97·9*	117·0*	99·0*	112·0*	95·0	90·3	90·7	89·8
Total N for each social class		228	218	198	165	689	625	361	384

* Indicates N for sub-group is less than 20.

TABLE 3B

MEAN SCORES ON NON-VERBAL TEST AT AGE 11 BY FAMILY SIZE AND BY SOCIAL CLASS (SEXES SEPARATELY).

	Social class (Sex)	I and II		III (Non-manual)		III (manual)		IV and V	
		B	G	B	G	B	G	B	G
Family size	1	108·2	107·8	107·0	105·7	100·4	102·2	98·5	102·8
	2	110·4	108·6	103·1	107·0	102·7	103·6	97·1	99·2
	3	107·9	108·6	104·8	106·7	100·0	101·8	98·8	98·9
	4	105·7	114·3	93·6	101·3	99·8	98·1	98·7	94·2
	5	103·1*	102·9*	101·0*	107·0*	95·9	98·7	92·6	96·9
	6+	105·4*	108·7*	102·0*	104·5*	95·3	90·4	92·3	90·9
Total N for each social class.		228	218	198	165	689	625	361	384

* Indicates N for sub-group is less than 20.

TABLE 4

REGRESSION COEFFICIENTS OF TEST SCORE ON FAMILY SIZE BY SOCIAL CLASS (SEXES GROUPED).

Test	Social Class	I and II	III (non-manual)	III (manual)	IV and V
English (12)		−1·091	−2·652	−2·565	−2·144
Non-Verbal (11)		−0·658	−1·545	−1·839	−1·612

These results seem to confirm Douglas's finding (1964) that the effect of family size on test score is less marked in social classes I and II. However, the small numbers in these classes make the regression coefficient liable to quite large sampling fluctuations.

One interesting finding is a consistent difference in regression coefficients between boys and girls on all tests. The test of significance involves assumptions which cannot strictly be met, and even if it used, the only sex difference in regression coefficients which reaches statistical significance ($P < \cdot 05$) is that on the Verbal Reasoning test at age 12. It is thus possible that these differences may be caused by sampling errors. Results for girls tested in 1958 as part of a study into menarche (Nisbet and Illsley, 1963) (Table 5 below) suggest that there may have been a similar effect in that age-group compared with the complete age-groups tested in other years. It is thus possible that differences in the treatment of boys and girls in large families could lead to sex differences in the relation between test score and family size. Douglas (1966) has also described results which show differences in measured intelligence apparently related to sex differences in the compositions and birth-orders of families.

IV.—RESULTS FROM SUCCESSIVE AGE-GROUPS IN ABERDEEN.

It has been possible to collect data on family size and intelligence test scores in a number of different studies carried out on comparable age-groups of Aberdeen children in 1949, 1950, 1956, 1958 and 1964. These successive checks on the relation between family size and intelligence test score suggest a possibility of detecting slowly changing social patterns, but the period covered is still too short to yield clear evidence. The scores from successive age-groups cannot be used to check any trend in the level of national intelligence as each test has been restandardised to a mean of 100.

Tables 5 and 6 summarize the changes in distribution of family size and in mean intelligence test scores for the various family sizes for each of the age-groups tested.

The differences in the distribution of family size in 1964, may be a result of an underestimation of the numbers of children in family sizes 1 and 2. The records used in this present investigation did not include data on family size for 303 children who had entered Aberdeen between 1962 and 1964, and this group is likely to contain a disproportionately high number of children from small middle-class families. This sampling difference should not affect the means or regression coefficients previously reported.

TABLE 5

PERCENTAGE DISTRIBUTION OF FAMILY SIZE, 1949–64.

	Year	1949	1950	1956	1958	1964	Douglas 1957* (National sample)
Family size ..	1	13	12	14	14	9	13·9
	2	25	28	30	34	31	32·7
	3	21	22	25	23	27	22·7
	4	16	14	14	13	17	14·3
	5	9	9	8	8	9	7·3
	6+	15	16	9	7	7	9·1
	(N)	(2561)	(2607)	(2248)	(1378)	(2868)	(8085)
Mean family size ..		3·4	3·4	3·0	2·9	3·1	3·0

* Reworked from data in *The Home and the School* (1964) by J. W. B. Douglas(page 168)

TABLE 6

MEAN SCORES ON VERBAL REASONING TESTS BY FAMILY SIZE ON DIFFERENT AGE-GROUPS BETWEEN 1949 and 1964.

	Year	1949	1950	1956	1958*	1964
Family size ..	1	104·9	105·7	104·5	105·3	105·2
	2	104·5	104·3	104·2	102·4	103·7
	3	102·3	101·4	98·8	99·7	101·6
	4	97·4	97·7	97·7	95·7	98·4
	5	94·3	93·6	94·9	91·7	96·2
	6+	91·6	91·5	90·4	91·4	92·8

* Girls only.

ACKNOWLEDGEMENTS.—The present study would not have been possible without the co-operation of the Director of Education for Aberdeen, Mr. J. R Clark, and his staff and also the help of the head teachers of all the primary schools in Aberdeen and many of their staff who administered the tests. Much of the statistical work involved in preparing the tables was carried out by Mrs. Jennifer Welsh, Research Assistant in this Department. Co-operation with members of the Aberdeen Medical Sociology Research Unit has allowed interchange of data between parallel studies of the same age-group.

V.—REFERENCES.

DOUGLAS, J W. B. (1964). *The Home and the School* London · MacGibbon and Kee.

DOUGLAS, J. W. B. (1966). The association of birth spacing and some other related factors with attainment. *Bull. Brit Psychol Soc.*, 63, page A8 (abstract).

HOROBIN, G. W., BYTHEWAY, W. R., and OLDMAN, D. J. Social differentiation of ability. (*In preparation.*)

NISBET, J. D. (1953) Family environment. *Eugenics Review*, 45, 31-40. (Reprinted in HALSEY, A. H., FLOUD, J., and ANDERSON, C. A., *Education, Economy and Society*, Free Press of Glencoe, New York, 1961).

NISBET, J. D , and ILLSLEY, R (1963) The influence of early puberty on test performance at age 11. *Brit J. Educ. Psychol.*, 33, 169-176.

(*Manuscript received 15th June, 1966*)

THE STABILITY OF THE WISC IN A PSYCHIATRIC GROUP

By R. K. TURNER, A. MATHEWS and S. RACHMAN
(*Psychology Dept., Institute of Psychiatry (Maudsley Hospital), London*)

SUMMARY. The re-test stability of the Wechsler Individual Scale for Children was assessed on a small group of psychiatric patients. The reliability coefficients obtained after a six-month interval were satisfactorily high but significant practice effect increases were detected on the Performance and Full-Scale scores. Some of the WISC subtests yielded very poor reliability coefficients. No evidence of an association between intelligence test scores and change in psychiatric condition was obtained.

I.—INTRODUCTION.

IT is often necessary to re-assess the intelligence of children who are receiving psychiatric care. The interpretation of the patient's initial test scores is facilitated by the substantial knowledge of the WISC which psychologists have now accumulated (see Littell, 1960; Buros, 1965). The interpretation of subsequent test scores is, however, complicated by the usual problems of 'practice effects' and the general retest reliability of the instrument. In his review of the literature, Littell (1960) pointed out that the effect of "repeated administrations of the WISC" is "especially in need of systematic investigation." Since Littell's article was published, some information has been forthcoming, but we have been unable to locate data for psychiatric groups.

Retest figures for normal S's were reported by Estes (1955) and by Gehman and Matyas (1956). Estes retested fourteen children from a low-economic group and eighteen from a high-economic group, after a three-year interval. Both groups showed WISC Full Scale increases—7·7 in the low-economic group and 4·9 in the other group. Gehman and Matyas retested sixty boys and girls (mean age 11 years) after a four-year interval and obtained a smaller increase—98·13 to 99·82. The Verbal Scale showed no change (96·9 to 96·7), but the Performance Scale increased from 99·87 to 103·23. None of these changes was statistically significant and the stability of the WISC was further evidenced by the retest correlations which ranged from ·74 to ·77.

Two reports deal with groups of backward children. Reger (1962) tested sixty-five educable retarded boys on three occasions separated by yearly intervals. The increases obtained on the F.S. (Full Scale) and P.S. (Performance Scale) scores at the second test were statistically significant. The means obtained on each occasion were as follows:

TABLE 1

REPEATED ADMINISTRATIONS OF THE WISC ON 65 RETARDED BOYS (REGER, 1962).
(Reproduced with the permission of the author and publishers of *Psycholological Reports*.)

	Test I	Test II	Test III
F.S.	71·6 (51·8)	76·2 (52·6)	78·8
V.S	74·0 (57·7)	75·4 (57·2)	76·4
P.S.	74·5 (54·3)	81·6 (56·2)	85·5

Throne, *et al.* (1962) retested thirty-nine retarded boys (aged 11—15 years) after only 3—4 months. No significant increases were obtained and their mean results are shown in Table 1 above, in parentheses. The retest correlations ranged between ·89 and ·95. The sub-tests also produced high retest correlations (·67 to ·84).

Coleman (1963) studied twenty-four boys (mean age 11·4 years) who were receiving attention for ' learning problems.' The F.S. correlation obtained after a fifteen-month interval (mean) was ·77 ; the V.S. correlation was ·62 and the P.S. ·81. The retest correlations for sub-test scores ranged between ·29 and ·68. The Picture Completion, Picture Arrangement and Comprehension tests all produced correlations lower than ·4.

The stability of intelligence test scores among twenty adolescent psychiatric patients (mean age 13·8 years) was studied by Hiler and Nesvig (1961) but, unfortunately, they used the adult Wechsler Test (W.B.II). The results of the retests, carried out after 2—3 months and again 12—14 months later, are nevertheless, pertinent. Significant increases were obtained on the P.S. (94·9 to 100·2) and F.S. (86·7 to 90·1) at the second test. No changes occurred between the second and third occasions of testing. A significant correlation (+0·55) between V.S increases and overall ratings of clinical improvement was also noted. The correlation between F.S. and clinical improvement was smaller (0·36) but still significant ; P.S. changes appeared to be unrelated to clinical improvement. These authors make the important point that the raised F.S. and P.S. means which emerged at the second test mask the fact that some patients actually did worse on this occasion. Finally, the pattern of retest increases recorded by Hiler and Nesvig (P.S. highest, then F.S., followed by V.S.) was also obtained in a study of adult patients carried out by Coons and Peacock (1959).

Several suggestions arise from the work mentioned above : the WISC is a stable instrument which commonly yields retest correlations of ·75 and higher. The only exception to date was obtained on the V.S. in an abnormal group of S's.

Full Scale and P.S. means tend to increase on retest while the V.S. appears to be insensitive to practice effects. The P.S. usually shows a larger increase (between 2—7 points) than the F.S. (between 1 and 5 points). These increases were not found in a severely backward group. Vernon (1960) among others, has noted that dull S's tend to show smaller practice effects than bright S's.

The greatest mean increase occurs on the second occasion of testing. A third re-test may or may not show a further increase in scores. This phenomenon has also been noted in other intelligence tests—increased practice produces diminishing returns (Vernon, 1960).

Even though mean increases are usually observed on retest, some S's obtain lower scores on the second occasion. Clinical improvement may be associated with F.S. and V.S. increases.

Finally, the retest reliability of some of the WISC sub-tests is in question.

II.—PROCEDURE.

Most of the children referred to the In- and Out-Patient clinics of the Children's Department at the Maudsley Hospital are tested by a psychologist on, or shortly after, admission. The psychological testing is carried out either by a permanent member of the staff of the Psychology Department, or by a student psychologist. Over a period of two years we have aimed to retest as many as possible after a six-month interval. The patients actually retested are a minority of the children seen, since in many cases it was impossible or impractical to arrange a retest (often the patient was discharged within the six months and could not be contacted, or the patient failed to keep the appointment arranged for retest). It was not possible to arrange that the same psychologist administered the tests on each occasion.

The Sample. At the end of the two-year period, retest data had been collected on thirty-seven children. Three patients were excluded since investigation indicated the presence of brain-damage There were eight girls and twenty-six boys in the remaining thirty-four cases. Visual inspection of the test scores for boys and girls suggested that marked differences existed between the sexes. The average scores of the girls were much lower than those of the boys, and their practice effects were particularly low. It was decided, therefore, that in view of a possible sex difference, the analysis would concern only the data on the boys. All subsequent data relates to test scores obtained by twenty-six boys who were retested after six-month interval on the WISC.

Age. The sample ranged from 7·7—15·0 years, with a mean of 12·1 years and a standard deviation of 2·0 years.

Diagnosis. The children were classified independently by two psychologists into four diagnostic categories, any disagreements being arbitrated by a third psychologist. The four categories, with the number of cases found in the sample, were as follows: predominantly neurotic personality disorders (six cases), predominantly anti-social conduct disorders (fourteen cases), habit disorders, e.g., enuresis (two cases), and mixed or unclassifiable (four cases).

Initial Level of I.Q. The sample was found to be of above average intelligence, possibly due to the presence of a referral bias. The mean scores of the three WISC scales, together with their standard deviations, were Full Scale I.Q. 110 (S.D.=12·39), Verbal Scale I.Q. 108 (S.D.=12·95) and Performance Scale I.Q. 111 (S.D.=13·68).

III.—RESULTS.

Test-Retest Reliability. As can be seen from Table 2, all three scales were found to be of adequate reliability. In order to facilitate the use of the data in predicting the score most likely to be obtained on retesting an individual patient, parts of the regression equation have been included in the last three columns of the table.

TABLE 2

W.I.S.C. TEST RETEST RELIABILITY.

	Initial Mean	Initial S.D.	Retest Mean	Retest S.D.	Test-Retest Correlation	B $\left(r \times \frac{S\,D.1}{S.D.2}\right)$	A $(\bar{X}_2 - B\bar{X}_1)$	Standard Error (s d.$_1\sqrt{1-r^2}$
Verbal Scale	107·77	12·81	108·5	12·20	0·81	0·77	25·71	7·21
Performance Scale	111·15	13·68	119·31	13·57	0·73	0·73	38·28	9·20
Full Scale	110·27	12·39	114·96	12·28	0·80	0·80	27·08	7·30

(Predicted score on retest=A+B (Initial score))

Practice Effects. As was expected, the mean scores on retest were higher than those obtained from initial testing, presumably due to practice effects. (The possibility that increases may be related to clinical improvement is examined later.) Comparisons were made between test-retest on all three scales using a correlated 't' test, and showed that there was no significant difference in the case of the Verbal Scale, while a highly significant difference was found in the case of the Performance Scale ($p < 0.001$) and a significant difference in the case of the Full Scale ($p < 0.01$). However, when individual results are examined, it can be seen that although the majority of subjects improved their scores, a small number obtained a poorer score on retest.

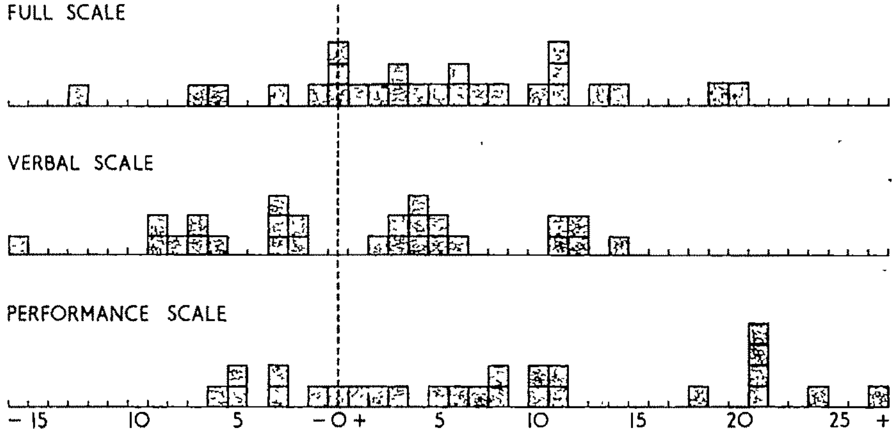


FIGURE I

Individual changes in Full Scale, Verbal Scale and Performance Scale scores on re-test after a six-month interval. It can be seen that some children do *worse* on the second occasion of testing.

Effect of Initial Score on Practice Effect. Product moment correlations were calculated between initial scores and the amount of increase on retest, with the expectation that the higher the initial I.Q., the greater should be the practice effect (Vernon, 1960). In fact, all the correlations proved to be negative; only the Verbal Scale reached an acceptable level of significance, indicating that subjects with lower initial scores tended to show the greatest practice effects. The correlations were, on the Full Scale, $r = -0.32$ (N.S.), on the Performance Scale, $r = -0.37$ (N.S) and for the Verbal Scale, $r = -0.46$ ($p < 0.05$).

It is possible that this finding is specific to the sample used, which was composed of a relatively large number of boys with high initial scores, and hence may show effects due to a regression to the mean. There does not seem any reason to suppose that the use of a psychiatric sample could account for this reversal of previous findings.

Subtest Reliability. Considerable variation was found in subtest reliabilities, and two (Picture Arrangement and Comprehension) failed to reach an acceptable level of significance (see Table 3).

TABLE 3

W.I.S.C. SUBTEST RELIABILITIES.

	Means (S.Ds.)				Reliability r 12
	Test I		Test II		
Comprehension	11·88	(3·63)	11·19	(4·96)	+0·33*
Similarities	10·96	(2·22)	11·42	(2·64)	+0·59
Arithmetic	10·65	(2·94)	11·65	(3·06)	+0·58
Vocabulary	12·08	(2·55)	11·5	(3·34)	+0·82
Digit Span	11·0	(2·97)	11·0	(2·90)	+0·75
Picture Completion	12·19	(3·43)	13·62	(2·80)	+0·54
Picture Arrangement	9·92	(2·27)	11·52	(3·00)	+0·11*
Block Design	12·92	(2·73)	14·0	(2·53)	+0·80
Object Assembly	12·58	(3·21)	13·88	(3·44)	+0·57
Coding	10·19	(2·45)	10·65	(2·77)	+0·61

* Not significant at the 5 per cent. level of confidence.

A Short Form of the WISC. Maxwell (1959) has published a shortened form of the WISC, based on a factor analysis of the original standardisation data. This Short Form consists of four subtests ; vocabulary and similarities (verbal factor) and block design and object assembly (visuo-spatial factor). As this form is used in clinical practice, the stability of the factors, together with their relation to the completed verbal and performance scales, respectively, was examined. The results, presented in Table 4, show the factors to have similar reliabilities to the complete scales.

TABLE 4

SHORT FORM RELIABILITY.

	Initial Mean	Initial S.D.	Retest Mean	Retest S D.	Test-Retest Correlation	Correlation with Complete Scale
' Verbal ' Factor	108·35	11·74	108·12	12·65	0·79	0·65
' Visuo-Spatial ' Factor	115·38	15·09	121·58	13·68	0·72	0·83

Score Change and Psychiatric State. A view commonly put forward is that improvement in a patient's psychiatric condition may be reflected in an increase of score on intelligence testing. While this may be true in some instances of severly disturbed patients, it is unclear whether this effect should be generally expected in clinical populations.

In order to rate pshyciatric changes, two psychologists read all reports made by psychiatrists or P.S Ws. in the period from initial testing to restest, and rated each patient on a five-point scale from Symptons Absent (5) to Definite Deterioration (1). This was done independently and without knowledge of WISC scores, any disagreement being arbitrated by a third psychologist.

Product moment correlations were performed between the rating of psychiatric change and each of the three WISC scales. The correlation of psychiatric change with increase in Full Scale was found to be –0·09, with Verbal Scale –0·18, and with Performance Scale +0·02. These results fail to support the view that intelligence test scores are associated with changes in the patients' clinical condition.

Reading Attainment. Results are also included of the Schonell Graded Word Reading Test which was also given on each occasion of testing. The results show a very satisfactory level of reliability (r=0·95), but also demonstrate a marked practice effect, with a mean increase of one year in reading age over the six-month period. The initial mean was 13·36 years with a standard deviation of 3·29, while the retest mean was 14·38 years, with a standard deviation of 3·2. This finding had been expected on the basis of previous studies on attainment tests (Curr and Gourlay, 1960).

IV.—DISCUSSION AND CONCLUSIONS.

The main conclusion to be drawn from the present study is that the results are surprisingly consistent with earlier findings obtained on non-psychiatric samples.

The retest stability of the WISC is reassuringly satisfactory. In regard to practice effects, the Verbal Scale appears to be insensitive to this type of inflation. The Performance Scale, however, is subject to a significant practice effect; in the present study, a mean increase of 8·16 I.Q. points. The Full Scale score is also subject to a practice effect, but to a lesser degree (mean increase of 4·7 I.Q. points) In practice, therefore, caution needs to be taken in assessing the possible meaning of WISC score increases as they are observed after a course of psychiatric treatment In the majority of cases, one can reasonably expect an increase on the Performance (++) and Full Scale (+) scores at retest. It must be remembered, however, that some children will do worse on retest.

Contrary to what one might expect, these gains and losses do not appear (at least in this sample) to be associated with changes in the child's psychological condition. One factor which is related to an increased Verbal Scale score is the patient's initial I.Q. Children with low scores can be expected to show larger practice gains.

The sub-test reliabilities show considerable variation and indicate, once again, the dangers involved in attaching exaggerated significance to sub-test patternings and 'profile analyses.' It is interesting to notice that, like Coleman (1963), we found the Picture Arrangement and Comprehension sub-tests to be particularly unreliable.

The results obtained with the Schonell reading test are consistent with other studies in showing a marked practice effect—coupled with a very high retest correlation. Like the intelligence test data, this emphasizes the need to know more about the stability of a test than is provided by a straightforward retest correlation.

Finally, we feel that the results of the present (and related) studies indicate that the WISC is sufficiently stable for repeated use. When interpreting test score changes, however, allowance must be made for the influence of practice effects on the Full and Performance Scales.

ACKNOWLEDGEMENTS.—We would like to acknowledge the help of all the students who have assisted in this study and, in particular, Miss P. Hay.

V.—REFERENCES.

BUROS, O. K. (1965). *The Sixth Mental Measurements Yearbook.* New Jersey : Gryphon Press.

COLEMAN, J. (1963). Stability of intelligence test scores in learning disorders. *J. Clin. Psychol.*, 19, 295-298.

COONS, W. H., and PEACOCK, E. P. (1959). Inter-examiner reliability of the WAIS with mental hospital patients. *O.P.A. Quart.* (July).

CURR, W., and GOURLAY, N. (1960). The effect of practice on performance in scholastic tests *Brit. J. Educ. Psychol.*, 30, 155-167.

ESTES, B. W. (1955). Influence of sociometric status on the WISC . *J. Consult. Psychol.*, 19, 22-226.

GEHMAN, I. H., and MATYAS, R. P. (1956). Stability of the WISC and Binet Tests. *J. Consult. Psychol.*, 20, 150-152.

HILER, E., and NESVIG, D. (1961). Changes in intellectual functions of children in a psychiatric hospital. *J. Consult. Psychol.*, 25, 288-292.

LITTELL, W. M. (1960). The Wechsler Intelligence Scale for Children : review of a decade of research. *Psychol. Bull.*, 57, 132-156.

MAXWELL, A. E. (1959). A factor analysis of the Wechsler Intelligence Scale for children *Brit. J. Educ. Psychol.*, 29, 237-241.

REGER, R. (1962). Repeated measurements with the WISC *Psychol. Reports.* 11, 418.,

THRONE, F. M., SCHULMAN, J. L., and KASPAR, J. C. (1962). Reliability and stability of the WISC for a group of mentally retarded boys. *Amer. J. Ment. Defic.*, 67, 455-457.

VERNON, P. (1960). *Intelligence and Attainment Tests.* London : Univ. of Lond. Press.

(*Manuscript received 24th June*, 1966)

SOME ASPECTS OF A STUDY OF THE GIFTED CHILD

BY K. LOVELL AND J. B. SHIELDS

(University of Leeds)

SUMMARY. This paper reports on selected aspects of a study of fifty 8 to 10-year-old pupils all of whom obtained a WISC Verbal Score of 140 or more. Four issues are dealt with in connection with these gifted pupils: personality, mathematical attainment, logical thought, and the relationship between scores obtained on creativity tests and those obtained on the WISC Scales and on tests of logical thinking. The findings suggest that their teachers rated their personalities very much as American teachers had rated Terman's comparable sample, and although performance on tests of mathematical attainment was high, these gifted pupils had the greatest difficulty with certain kinds of series and numerical analogies. Only 10 per cent. of the responses were at the level of formal thought on tasks of logical thinking, while a principal components analysis showed that tests of creativity do not measure intellectual functions entirely independent of those measured by the WISC Scales and tests of logical thought.

I.—INTRODUCTION.

A STUDY has recently been undertaken to obtain current information about the intellectual abilities, language and the personality traits of the gifted child in England ; the gifted child being defined as one who obtained a score of 140 or more on the WISC Verbal Scale. This paper reports on selected aspects of this study. There are three major studies in this field. Terman (1926) selected children if they obtained an I.Q. of 140 or more at 11 years of age or under, and a score adjusted to 140 if over 11 years, using the Stanford-Binet Intelligence Test. He and his associates inquired into the abilities, interests and background of their subjects. By means of a medical examination, a study of school records and tests, home visits, and from statements by parents and teachers, detailed information about each individual was acquired. Parents and teachers were asked to rate the subjects on a list of twenty-five personality traits. Parkyn's (1945) study was designed to provide information about highly intelligent children in New Zealand, while in the second half of this book he considers the proper education of the pupils. Of the three groups he studied, only one is in any sense comparable with Terman's sample. This was made up of fifty pupils who obtained a score on the Revised Stanford-Binet Intelligence Test (1937) which placed them in the top 5 per cent. of their age group. The majority of these subjects were 12 years of age. The third study was carried out by Burt. In 1915, amongst many other studies, he began the practice of considering the abilities and the performance of the top 3 per cent. of candidates sitting the then London County Council Junior County Scholarship examination. Information was obtained on school performances, on leisure interests, on the occupation of the parents, and on the home background of the children. Burt (1962a) reported that his findings were in general agreement with those of Terman.

In recent years Guilford's convergency-divergency dimension* led Getzels and Jackson (1962) to see a possible relationship between this dimension and the problem of the identification of the gifted. Their work was critically reviewed by Burt (1962b) and Vernon (1964). Although many studies have been carried out into the general question of creativity (e.g., Torrance, 1962; Yamamoto 1964a, 1964b, 1965; Clark, Veldman and Thorpe, 1965) none has been concerned with pupils comparable with those in Terman's sample. Another problem of considerable importance is whether or not gifted children move into Piaget's stage of formal thought much earlier than pupils of average ability-as measured by I.Q. tests. The research reported here deals with just four issues·

---

* The recent work of Harris and Liba has thrown some doubt on the validity of Guilford's model of the structure of the intellect. See Harris, C. W. and Liba, M. R. (1965) : *Component, Image and Factor Analysis and Tests of Intellect and of Motor Performance.* Co-operative Research Project No. 3-192-64. University of Wisconsin.

(*a*) personality as assessed by ratings;

(*b*) the level of mathematical attainment;

(*c*) the growth of logical thought;

(*d*) the relationship between scores obtained on creativity tests and those obtained on the WISC Scales and on tests of logical thinking.

II.—PLAN OF THE STUDY.

(i) *Subjects.* All the children were drawn from two large cities in the North of England. The schools were asked to nominate children of outstanding intellectual ability. Over one-hundred pupils, thus nominated, were tested using the WISC with the alterations suggested by the British Psychological Society Verbal, Performance and Full Scale I.Qs. were obtained. The first fifty pupils who obtained an I.Q. of 140 or more on the Verbal Scale formed the selected group; the numbers of children with I.Qs. in the ranges 140-144, 145-149, 150-154, 155+ being respectively 30, 12, 4 and 4. Thirty-five boys and fifteen girls were chosen, while forty of the fifty pupils were found in maintained schools. Twelve pupils were between 8·5 and 9·4 years of age, twenty-one were between 9·5 and 10·4 years, and seventeen were between 10·5 and 11·7 years old. Using the Registrar-General's Classification of Occupations (1960), it was found that nine pupils came from homes in Social Class 1, twenty-two in Social Class 2, fifteen in Social Class 3, and four in Social Class 4.

(ii) *Tests used.*

1. WISC Verbal Scale.
2. WISC Performance Scale.
3. WISC Full Scale.

Divergent Thinking (Creativity) Tests*.

4. Hidden Shapes (as used by Lovell, 1954, and similar to that used by Getzels and Jackson).
5. Word Association. Subjects were required to give as many different meanings as possible to twenty-five fairly common words; e.g., arm, cap.
6. Uses for Things. Subjects had to write down as many ***different*** uses as possible for five common objects.
7. Fables. Four incomplete fables were given. Subjects were asked to compose three different endings for each fable: a moralistic, a humorous and a sad ending.
8. Make-up Problems. Three paragraphs were given each containing many numerical statements describing situations well within the experience of the subjects. They were asked to make up as many problems as they could using the information given.

---

*Details of these tests and the procedures used in scoring may be obtained from the authors.

Tests of Logical Thinking.

9. Equilibrium in the Balance (Inhelder and Piaget, 1958).
10. Combination of Colourless Liquids (Inhelder and Piaget, 1958).
11. The Oscillation of a Pendulum (Inhelder and Piaget, 1958).
12. Concept of Volume (Lovell and Ogilvie, 1961).

Mathematics Tests.

13. N.F.E.R. Number Concept Test (7+) N.S.44.
14. N.F.E.R. Concept Arithmetic Test (10+) N.S. 54.
15. N.F.E.R. Concept Test ' N.'
16. N.F.E.R. Mathematical Insight Test (14+), N.S. 63.
17. A Series and Numerical Analogies Test.
18. Vernon's Graded Arithmetic—Mathematics Test.

All the testing was performed in the pupils'school and was on an individual basis, except in the case of the Mathematics tests where up to three children were tested at the same time. In addition, a Terman Personality Trait Rating Form was completed by a teacher for each child for each of twenty-five personality traits. Thirteen grades were used for each trait.

III.—RESULTS.

(a) *Personality Ratings.*

They varied from a mean rating of 2·4 for General Intelligence to 6·14 for Freedom from Vanity and Egotism. The means of all the ratings were less than seven. It is likely that there is some ' halo effect ' operating, but it may be noted that Terman felt it safe to assume that such an influence was also operating in the ratings of his control group. In the opinion of the teachers, the gifted child is ' above average ' on all the personality traits rated. These gifted pupils were rated as outstandingly high in General Intelligence and Desire to Know ; very high in Originality, Desire to Excel, Truthfulness, Commonsense, Will Power and Perseverance, also Conscientiousness ; rather high in Prudence and Forethought, Self-Confidence and Sense of Humour. In a few traits they were rated close to average, the poorest rating being in respect of Freedom from Vanity and Egotism. There were a few sex differences. Boys were rated somewhat higher on Sensitiveness to Approval or Disapproval, Amount of Physical Energy, Fondness for Large Groups and Mechanical Ingenuity ; whereas girls were rated rather higher for Appreciation of Beauty, Musical Appreciation, and Freedom from Vanity and Egotism. The mean ratings given by British teachers on the traits were very close to those given by American teachers to Terman's sample some forty years ago, while when the traits were placed in rank order for the present study and for Terman's group, Spearman's Rank Difference Correlation Coefficient was +·90.

In addition, the teachers' ratings of the subjects for each of the twenty-five traits were ranked, and the rankings transformed to a normal distribution. Product moment intercorrelations for the twenty-five variables were calculated, and a Principal Components Analysis together with a rotation of the principal axes using the Varimax method was carried out on the Leeds University Computer. Six components were extracted (all eigen values greater than 1) ; the percentage variance contributed by these was respectively 25·0, 12·3, 10·3, 8·3, 6·9, 5·3. Table 1 shows the principal component axes and the rotated axes, with loadings which did not reach the 5 per cent. level of significance omitted.

TABLE 1

PERSONALITY TRAITS : UNROTATED AND ROTATED AXES.

	Principal Component Axes						Rotatee Axes					
	I	II	III	IV	V	VI	I	II	III	IV	V	VI
Health	52						56					
Amount Phys/Energy	55	56					81					
Prudence and Forethought			52		–53						84	
Self-confidence			65									
Willpower/Perseverance	56		61								73	
Musical Appreciation				50		54						76
Appreciation of Beauty		–51						61				
Sense of Humour	61							54				
Cheerful/Optimism	46	50					64					
Permanency of Moods								65				
Fondness/Large/Groups	54	69					85					
Leadership		66					76					
Popularity with other Children	66						60					
Sensitive to Approval/Disapproval												70
Desire to Excel	49			45			47				55	
Freedom from Van/Egotism		–53			–50			51	57			
Sympathy/Tenderness	50		–62						73			
Generosity/Unselfishness	64		–48						70			
Conscientiousness	58					–45			57		54	
Truthfulness	57								63			
Mechanical Ingenuity	51			–44				71				
Desire to Know	46				59					91		
Originality	53				57					85		
Common Sense	72							67				
General Intelligence	62									69		

(Decimal points omitted)

(*b*) *Mathematical Attainment.*

On Mathematics Tests 13, 14, 15, 17 and 18, the scores were high as expected. For example, on Vernon's Graded Arithmetic Mathematics Test the mean achievement age was 3 years 7 months in advance of the mean chronological age. But twelve pupils were from 4 to 5 years in advance of their chronological age level, three were 5 years in advance with one pupil actually 5 years 9 months in advance. This subject also scored very highly on Tests 13, 14, 15 and 16. However, it was extremely interesting to note the great difficulty which these gifted pupils had with certain of the problems involving Series and Numerical Analogies. Whereas forty-eight subjects could continue the series 48, 24, 12, 6 . . . , only eleven could continue the series, 16, 24, 36, 54 . . . , and only five could answer the questions ' 3 is to 7 as 9 is to . . . ,' and ' 2 is to . . . is to 8 as 3 is to 9 is to ... '.

(*c*) *The Growth of Logical Thinking.*

Responses to the Balance, Chemicals and Pendulum Experiments were classed as at stage 1B/2A (Preoperational) ; stages 2A, 2AB, 2B, 2B/3A (Concrete operational) ; and stages 3A, 3AB, 3B (Formal operational). The number of responses at each of the stages were respectively 3, 22, 41, 49, 20 ; 11, 3, 1. It is clear that only 15/150 responses were at the stage of formal operational thought and eleven of these fifteen replies were at stage 3A, i.e., at the first stage of formal thought. The one pupil who attained stage 3B reached this on the Chemicals experiment.

Almost all the pupils understood internal volume, occupied volume and simple displacement. But only thirty-six understood that the amounts of water displaced by a single cube when immersed into a full pint can and a full gallon can were the same; while twenty-six grasped that the amounts of water displaced by two cubes of the same size, but different weights were equal. Overall, only twenty-four subjects (48 per cent.) answered all the questions successfully, while the proportion of the younger subjects (42 per cent.) is not greatly different from the proportion of the older pupils (53 per cent.) passing all the tests.

(d) *The Relationship between Scores obtained on Creativity Tests and those obtained on the WISC Scales and on Tests of Logical Thinking.*

Tests 12, 13 and 16 were omitted from the analysis as the raw scores were markedly skewed. In the case of the other tests the raw scores were recast in the form of a normal distribution so that the scores ranged from +5 to –5. To get a set of normal scores for the WISC Full Scale, the raw scores on both the Verbal and Performance Scales were recast into normal form, added, and the total score so obtained again recast into a distribution ranging from +5 to –5. In the case of tests 9, 10, 11, numerical values were awarded to the stages (Stage 1B/2A, as 1, Stage 2A as 2, etc.) and these values recast in normal form.

Two further measures were obtained, namely, scores for the Creativity Tests summed, and scores for the tests of Logical Thinking summed. Product moment intercorrelation coefficients for the seventeen measures were calculated, and a Principal Components Analysis and rotation of the principal axes by the Varimax method were carried out. Six components were extracted; the percentage variance contributed by these was respectively 39·3, 12·4, 9·7, 7·1, 6·3, 5·6. Table 2 shows the principal component axes and rotated axes with loadings which did not reach the 5 per cent. level of significance omitted.

TABLE 2

COGNITIVE TASKS: UNROTATED AND ROTATED AXES.

Test	Principal Component Axes						Rotated Axes					
	I	II	III	IV	V	VI	I	II	III	IV	V	VI
WISC Verbal Scale	77		–39				77			54		
WISC Performance Scale	54	–46	–54				86					
WISC Full Scale	76		–53				88					
Hidden Shapes	42					57						86
Word Association	64	44		–34					76			
Uses for Things	41	45		–39					82			
Fables		53			61				36		83	
Make-up Problems	37	40		37	–45					80		
Creativity Summed Scores	73	56							82			
Balance	37	–50	45			–37		83				
Colourless Liquids	61		37					72				
Pendulum	72		38					60			42	40
Logical: Summed Scores	73	–40	50					93				
Mathematics 10+	79							37		51		
Mathematics 'N'	71			35			45	34		49	43	
Vernon's Graded A/M	67			44			39			35	73	
Series and Numerical Analysis	72						39		45	42		

(Decimal points omitted)

IV.—DISCUSSION.

It has already been pointed out that the gifted child in Britain seems to be above average on all the traits listed in Terman's Rating Scale. But equally interesting is the fact that despite changes in education and life generaly, both in the U.S.A. and in the U.K., between the 1920's and 1960's, the gifted child seems to his teacher to be much the same kind of person.

The Principal Component Analysis of the personality trait ratings, after rotation of axes by the Varimax method, suggests six clear dimensions. Number one indicates Gregariousness, Leadership, Health, Physical Energy and a Cheerful Outlook associated with a measure of Popularity with other children and a Desire to Excel. The second dimension indicates Practical Inventiveness accompanied by Common Sense, an Eye for Beauty, a Stable Temperament, Sense of Humour and Sense of Modesty. Moral traits load the third dimension, although it should be stressed that Sympathy, Generosity, Truthfulness and Conscientiousness seem linked with a lack of Self Confidence in these very able pupils. Dimension four suggests a combination of intelligence and motivation, (Desire to Know ·91, General Intelligence ·69), while the fifth and sixth indicate respectively, a volitional dimension (Prudence ·84, Will Power and Perseverance ·73), and the possibility of Musical Appreciation linked with Sensitivity to Approval and Disapproval.

The mathematical attainment of gifted children is much what was expected as far as performance on most of the tests was concerned It is, however, important to note the great difficulty that they experienced in certain types of series and numerical analogies that demand for their solution, the schema of proportion. This schema is not available to the child until he has reached the stage of formal operational thought (Lovell, 1961, 1966; Lunzer, 1965; Lunzer and Pumfrey, 1966) and it does appear that few of these gifted pupils have attained this, at least in the areas in which they were questioned. The responses that they gave to the tasks involving logical thought clearly confirm this. Only 10 per cent. of these could be classified as at the stage of formal thought, and most of these were no higher than stage 3A. Of the five pupils who successfully answered the question ' 3 is to 7 as 9 is to . . . ', two were at the stage of formal thought in the Balance experiment and one was at the borderline stage of 2B/3A. This finding once again brings out the point that formal thinking is not available to the child in all situations at first

On the other hand, the responses to the tasks involving the concept of volume shows that a far higher proportion of gifted pupils could answer the questions correctly than in the case of a cross section of junior school pupils (Lovell and Ogilvie, 1961). One might suggest, therefore, that while these pupils do not generally attain the level of formal operational thought until 11 years of age at least, it seems that the flexibility of their first order operational schemas is much greater than in the case of ordinary pupils.

The Principal Components Analysis of the scores for the cognitive tasks clearly shows (Table 2) the presence of a large general component, and a bipolar component separating the tests of logical thinking from the more ' learned ' abilities underlying performance in the WISC. On the rotated axes the dimension of logical task performance is clearly seen; the summed score of the logical tasks had a loading as high as ·93 on it.

Table 2 also indicates the large general component running through the divergent thinking and other conitive tasks even within this highly selected group; indeed, it accounts for one-half of the identified variance. The first bipolar component separates the tests of divergent thinking (excluding Hidden Shapes) from the WISC Performance Scale. When the axes are rotated, the first dimension clearly indicates an ability measured by the WISC Scales or $g+v:ed$ ability; the second reflects a capacity to think logically in the Piaget type situation, while the remaining dimensions suggests that divergent thinking cannot be accounted for by one dimension; rather the able pupil is ' creative ' to different degrees according to the task that is set him. While the verbal creativity tests load the third dimension, the ability to make up numerical problems is clearly linked with the ability to work the mathematics tests. This study supports Burt when he wrote of divergent thinking tests (Burt, 1962, p. 295) " . . . that these new tests have succeeded in eliciting supplementary activites that are rarely tapped by the usual brands of intelligence tests." On the other hand, the present evidence is clear that a great part of the identified variance of these tests is accounted for by a central intellective component, which is common to a conventional test of intelligence, to tasks involving logical thinking, and to tests of mathematical attainment. Moreover, these ' supplementary activities ' seem to fragment.

Finally, the normalised scores on the seventeen cognitive variables were ranked and combined with twenty-one of the personality trait ratings (those for Health, Permanency of Moods, Desire to Excel, and Freedom from Vanity and Egotism being omitted) to yield thirty-eight variables in all. Product moment intercorrelations were calculated, a Principal Components Analysis was carried out and the principal components axes rotated by the Varimax method. Six components were obtained but they largely repeated the components found in the earlier analysis. However, one point is of interest. The teachers' ratings of Originality load the same rotated axis (·84) as does the Logical Thinking Summed Scores (·63), yet Originality has a zero loading on the axis which reflects the ability to answer the tests of divergent thinking (Creativity Summed Scores ·82). It would seem that using Terman's Personality Rating Scale, teachers think of Originality in these pupils in terms more of reasoning ability than they are in terms of inventiveness or unusual ideas; or it may be that ' creativity ' measured by these tests is too specific and at times too trivial to be noticed by teachers.

V.—REFERENCES.

BURT, C. (1962a). The gifted child. In *The Year Book of Education*. London . Evans, pp. 24-58.

BURT, C. (1962b). Critical notice: creativity and intelligence, by J. W. Getzels and P. W. Jackson. *Brit. J Educ. Psychol.*, 32, 292-98.

CLARK, C. M., VELDMAN, D. J., and THORPE, J. S. (1965). Convergent and divergent thinking abilities of talented adolescents. *J. Educ. Psychol.*, 56, 157-163.

GETZELS, J. W., and JACKSON, P. N. (1962). *Creativity and Intelligence.* London: Wiley.

INHELDER, B., and PIAGET, J. (1958). *The Growth of Logical Thinking from Childhood to Adolescence.* London: Routledge and Kegan Paul.

LOVELL, K. (1954). A study of the problem of intellectual deterioriation in adolescents and young adults. *Brit. J. Psychol.*, 49, 199-210.

LOVELL, K. (1961). A follow-up study of Inhelder and Piaget's ' The Growth of Logical Thinking.' *Brit. J. Psychol.*, 52, 143-153.

LOVELL, K. (1966). The growth of the schema of proportion. *Mathematics Teaching*, 37, 5-9.

LOVELL, K., and OGILVIE, E. (1961). The growth of the concept of volume in junior school children. *J. Child Psychol. and Psychiat.*, 2, 118-26.

LUNZER, E. A. (1965). Problems of formal reasoning in test situations. In European research in cognitive developments. *Monogr. Soc. Res. Child Developm.*, 30, No. 2, 19-46.

LUNZER, E. A., and PUMFREY, P. D. (1966). Understanding proportionality. *Mathematics Teaching*, 34, 7-13.

PARKYN, G. W. (1948). *Children of High Intelligence.* London: Oxford University Press.

TERMAN, L. M. (1926). *Genetic Studies of Genius, Vol. 1—Mental and Physical Traits of a Thousand Gifted Children.* Stanford University Press.

THE THEORY OF FLUID AND CRYSTALLIZED GENERAL INTELLIGENCE CHECKED AT THE 5-6 YEAR-OLD LEVEL

BY RAYMOND B. CATTELL

(*University of Illinois*)

I.—THE HYPOTHESIS TO BE TESTED AND ITS BACKGROUND IN THEORY.

ACCORDING to the theory of fluid and crystallized general ability, there is not *one* general ability factor, as originally propounded by Spearman (1904) and supported by Thurstone (1938), but *two*. It states that these two broad* factors are distinguishable by one, called crystallized intelligence, g_c, loading most heavily the culturally acquired judgmental skills, while the other, called fluid ability, g_f, is found loading insightful performances in which individual differences in learning experience play little part.

That earlier researches did not discover this double structure seems partly due to the comparative crudity of early factor analytic techniques, e.g., in relation to estimation of communality and factor number, but especially to the highly prevalent defect of design which omitted the inclusion of *sufficient 'hyperplane stuff' to permit reliable rotation of any factor absolutely general to the whole ability domain*. When these circumstances were remedied, by providing a large enough 'ground' for factors spanning most cognitive performances to be rotated as discernible patterns, the resolution was plainly different. Now functionally independent, but appreciably correlated ($r = +0{\cdot}47$), g_f and g_c factors appeared (Cattell, 1963), where only one g had been posited. This finding with 12–13-year-old children has been confirmed, on a larger scale, by Horn's study with adults (1964).

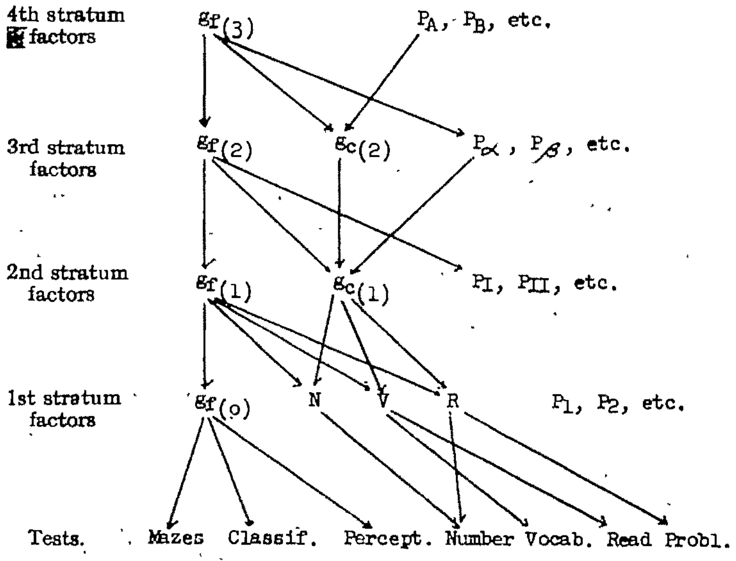
Demonstration of the *factorial separability* of the two concepts has been only a beginning, though a necessary one, for checking and extending the remaining parts of the theory which have to do with the differences of g_c and g_f in (*a*) age change plots, (*b*) population scatter of I.Qs., (*c*) relations to brain injury, (*d*) size of nature-nurture ratios, (e) expression in culture-fair as opposed to traditional intelligence tests, etc. (Cattell, 1963 ; Cattell and Horn, 1966 ; Hebb, 1942 ; Horn, 1966), and (*f*) a development theory of the generation of crystallized general ability by the joint action of fluid ability and experience-plus-motivation (whereby the fluid ability level over the preceding years would determine, in part, the crystallized ability this year). With these further implications, the present research is not directly concerned. It invokes them only insofar as they have implications for the hypothesis and design of experiments extending investigation of g_c and g_f factor structure to the early age of 5–6 years, with which this experiment is specifically concerned. However, Diagram 1 is a brief schematic summary of what can be read in greater detail elsewhere (Cattell, 1963 ; Horn, 1964).

---

* As argued elsewhere (Cattell, 1966), 'broad' is a better antithesis to specific than is 'general.' In nature, no factor is general to all variables, nor does even the matrix in a well-designed single experiment permit a general factor—else it becomes rotationally indeterminate. The g_c and g_f factors, according to Horn's large experiment (1964), although not absolutely general even to the cognitive domain, are quite the broadest factors. However, two other factors approach them in breadth—visualization and speed. The former, by breadth of cognitive effect, has *some* claim to being called a third form of intelligence. If this were granted, however, there are probably other broad sensory structuring functions—auditory, kinesthetic, olfactory—which could claim the same. One may perhaps anticipate that these will form a collateral set of *relatively* broad, but not general, 'sensory and relational resolution powers' below general intelligence.

DIAGRAM 1

Causal Sequences in Generation of Primary Abilities.



(The P's are personality factors, described later).

The theory of fluid ability is that it produces individual differences mostly in perceiving relations in (*a*) material equally new to all and (*b*) material equally 'overleant' for all. Where the group has persons with widely varying degrees of familiarity with the material, less of the variance will be due to g_f and more to g_c. However, unless this length of experience and familarity covaries for a wide array of intellectual performances, this variance will show more in specifics than in g_c. Since more school-type performances are new to 5-year-olds than say, 15-year-olds, we should expect that g_f would affect more types of performance, e.g., kinds of school, subject performances, in 5 than in 15-year-olds (Hypothesis 1). However, since the percentage of *variation* in 'time spent at school' is higher for 5 than 15-year-olds, we should expect experience to play a greater rolê than do the differences in g_f in building up g_c. The actual communality of those performances which *do* form a g_c could thus be very high. Consequently, g_c should be narrower than at later school ages and g_c and g_f should be less substantially positively correlated than in older persons, since differences of experience have had more to say than differences of fluid ability (Hypothesis 2). Since the uniformity of learning experience and the power of 'aids' (Cattell, 1963) is likely to prevail in *special areas*, e.g., mastering numbers, learning to read, more than across the *whole* scholastic and cultural field, at this age (compared to later ages, when aids have sunk into the background), we might expect primary ability factors (or, at least, numerical ability and reading ability) to have greater absolute loading on performances of this kind at 5 than later, and to be more separate. For example, reading speed, vocabulary size, and precision of spelling would be larger and more inter-correlated than in adults (Hypothesis 3). The crystallized ability factor should show up as a second order factor across these primaries (as it does at all ages), but because of the greater chances of uneven learning experiences in the different primary fields, it should not show as high a communality, i.e., should leave more specific variance *in the primaries*, than later (Hypothesis 4), i.e., the common variance in crystallized abilities should show up more in the primaries and less in g_c.

II.—DESIGN OF THE EXPERIMENT.

The choice of variables was carefully made to include (*a*) ability measures—twenty in all—covering some which would definitely be considered fluid ability, some theoretically largely involved in crystallized ability, and some which would be expected to load both, and (*b*) personality measures which could be expected to be uncorrelated with ability. These last are also chosen to ' spread out ' well in personality space, i.e., it is necessary that there be definite evidence that several independent dimensions will be involved. The provision in the hyperplane stuff of enough independent dimensions is important if we are to insure the condition, at the anticipated higher order factor resolutions, that enough distinct non-ability variables will exist to give dependable rotation to the broad ability factors expected to emerge at the higher order. To provide this condition, we took material demonstrated to represent thirteen distinct personality factors in the questionnaire medium, as recently discovered by Cattell and Peterson (1958).

The resulting list of variables is shown in Table 1. The last twenty-six variables are quite short questionnaire scales or ' parcels,' each containing one to four items found most loaded on the Cattell-Peterson nursery school factors. Each such factor is represented by *two* parcels, though we did not expect, nor is it necessary, that they be represented with high variance (for in at least four cases, the ' scale ' was only a single item in length).

TABLE 1

LIST OF COGNITIVE AND PERSONALITY MEASURES APPLIED IN THE EXPERIMENT.

No. in Matrix	Cognitive Ability Variables.
1	Substitution. No. completed.
2	Substituion. Inverse No. of errors
3	Classification (non-verbal).
4	Mazes.
5	Selection of Named Objects.
6	Following Directions.
7	Detecting Wrong Pictures.
8	Riddles.
9	Similarities (non-verbal).
10	Cal. Ment. Mat. (Language).
11	Cal. Ment. Mat. (non-lang.)
12	Reading Test. Word Mastery.
13	Reading Test. Sentences.
14	Reading Test. Information Gained.
15	Reading Test. Following Math.
16	Handling numbers.
17	Copying.
18	(Ach.) Test of Readiness for Reading Work.
19	(Ach.) Test of Readiness for Number Work.
20	Drawing a Man (Scored for ' Intelligence.')
	* IPAT Culture Fair Scale 1 (1-9).

TABLE 1—continued.

No. in Matrix	Personality Variables*	PSPQ Items (Indexed as in Cattell and Dreger)
21	Cattell-Peterson No. 1 (a) (G$_{(a)}$).	(1)
22	C-P No. 1b (G$_{(b)}$)	(11, –42).
23	C-P No. 2a (Fa)	(–10, –31, 41, 57)
24	C–P No. 2b (Fb)	(109, 177, 179, –196)
25	C–P No. 3a (Aa)	(21, –33, 52)
26	C–P No. 3b (Ab)	(–40, 134)
27	C–P No. 4a (Da)	(4)
28	C–P No. 4b (Db)	(13, 17)
29	C–P No. 5a (Ia)	(8, –27, –39)
30	C–P No. 5b (Ib)	(–140, –165)
31	C–P No. 6a (Ea)	(–6, –15, 46, –156)
32	C–P No. 6b (Eb)	(69, 111, 184)
33	C–P No. 7a (Oa)	(9, 25)
34	C–P No. 7b (Ob)	(–28, 51)
35	C–P No. 8a (Ma)	(–7, –26, 38, –92)
36	C–P No. 8b (Mb)	(–47, –74, 101, –127)
37	C–P No. 9a (Ca)	(12, –23, 30, –45)
38	C–P No. 9b (Cb)	(67, 192)
39	C–P No. 10a (Ha)	(–16, –29, 55)
40	C–P No. 10b (Hb)	(–54, –178)
41	C–P No. 11a (Ma)	(–14)
42	C–P No. 11b (Mb)	(5, 44)
43	C–P No. 12a (Q_2a)	(3)
44	C–P No. 12b (Q_2b)	(36, –56)
45	C–P No. 13a (Q_4a)	(50)
46	C–P No. 13b (Q_4b)	(–18, –20)

The ability tests consisted first of the eight subtests (one scored two ways) from the IPAT Culture Fair Intelligence Scale 1. As the Handbook (Cattell, 1950) explains, although the test is in the Culture Fair series, only a core of four subtests is considered substantially culture fair, i.e., of a type to measure fluid ability. (For in many practical situations, users wish to have a longer test where culture-fairness is no longer completely necessary.) To represent traditional tests, the California Mental Maturity (Intelligence) Test was included, but separating the score for the non-language part from the language score because, by content, the former could approach a culture fair subtest. After that, the ability measures are meant largely to represent reading and number performance, in various manifestations, as well as such skills as copying and drawing (the latter, however, scored for ' intelligence ').

The choice of subjects necessarily had to be nursery school children, with a sprinkling of first grade children. Some slight social selection is implied here, which should be kept in mind in later comparisons, in that it might produce a slightly early rise of crystallized ability structures. The age range was kept minimal to avoid spurious correlations of g_f and g_c (since both increase sharply with age). The characteristics of the subject sample are given in Table 2.

The testing was carried out by Dr. Larry Killian and Dr. John R. Nesselroade, to whose careful work we are much indebted. The intelligence test was given to groups of three and four children at a time and covered two sessions. The personality questionnaire items were administered similarly over three sessions. Voice-recorded instructions were used throughout the course of the

TABLE 2

PROPERTIES OF SUBJECTS.

Male=57 Female=57 Total=114

Mean Age=79·24 months (6 years 6 months)

S.D. of Age=5·17 „ „

test, and short rest periods were given about every ten minutes. The reading and number tests, the drawing of a man, and the California Test were given as a series of tests by the school to all children, and administered by the teachers of the different classrooms. Attention and interest by the children was maintained at a good level and we have no reason to believe that any major effects arose from motivation differences. The author is greatly indebted to the University of Illinois Nursery School and the Lottie Switzer and Garden Hills Schools, Champaign, for good advice and patient co-operation. The testing extended over the year 1963-64, and the sample in Table 2 represents those remaining when children with an incomplete test score were dropped.

III.—ANALYSIS OF THE DATA.

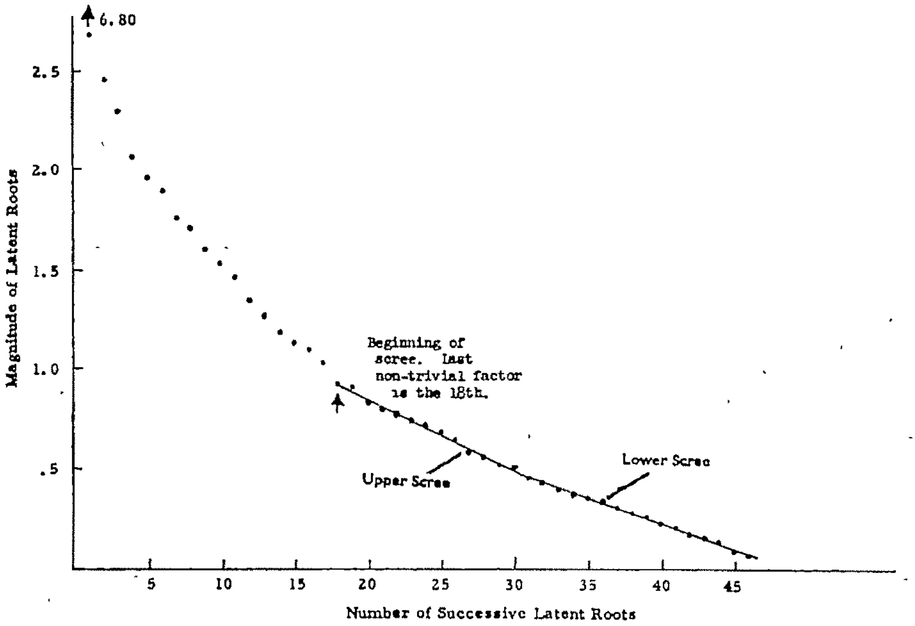
The testing of hypotheses calls for a first order factor analysis, which is then to be pursued into as many higher strata analyses as may be necessary to get complete resolution of the factors affecting the ability performance, just as has been attempted, up to a point, by three earlier studies (Cattell, 1963; Horn, 1964; Thurstone, 1938). For the main theory, as it affects cross-sectional studies, expresses itself as a model of higher order relationships.

For the first order, the product-moment R matrix among the forty-six variables was subjected to a principal axis analysis, with unities in the diagonal carried to forty-six factors. Application of the Scree test (Diagram 2) clearly indicated eighteen factors (the Kaiser-Guttman criterion suggested seventeen). Six iterations yielded communality convergence at eighteen factors. The application of the Maxplane (Eber, 1966) automatic topological program for simple structure was followed by twenty-one blind rotations by Rotoplot (Cattell and Foster, 1963). The successive hyperplane counts ($\pm$·10) were plotted on a curve until an unimprovable plateau was reached at 50 per cent. in the $\pm$·05 and 73 per cent. in the $\pm$·10 hyperplane. (This is in the V_{rs}, i.e., reference vector structure: the corresponding and scarcely different values in the V_{fp}, i.e., factor pattern, can be seen in Table 3.)

The V_o (unrotated), L (transformation), and full V_{fp} (factor pattern) matrices may be obtained from ADI under No. A111. Since we are not concerned here with the personality factors *per se* (which exist only for hyperplane stuff) and only one of them, No. 3 (No. 7 in the V_{fp}) has noticeable loading on cognitive variables (namely, on the assiduous 'substitution coding' performances, Nos. 1 and 2 in Table 1) nothing need be said about them, except that they do confirm the markers for the Cattell-Peterson personality factors listed on the left of Table 3. (Factor 18, though wholly of a personality nature, is an extra and seems to represent something at the personality second order). The personality factor structure as such is, in any case, set out fully elsewhere (Cattell and Dreger, 1967) so that only the first four primaries—the cognitive factors, arranged to come first—are shown in Table 3. (Drawing ability, as well as substitution skill, were exceptions among abilities in having loading on personality factors, namely, CP6; —0·35.)

DIAGRAM 2

Decision on Number of Factors by Scree Test: First Order



Inspection of the full eighteen factor V_{fp} matrix (ADI No. A111) shows immediately that the design has fulfilled its expectations. The thirteen Cattell-Peterson personality factors have replicated themselves and are as fully determinate, by the two markers for each, as one could expect from the short scales. They separate clearly as a spectrum of primary personality factors approximately orthogonal, but actually precisely angularly oriented among themselves. They give the required distinct background to the four factors, set out in Table 3, which together handle all the cognitive performances. Among the latter, the presence of emphatic number facility (N) and verbal facility (V, though much concerned with reading) factors is at once evident. Interpretation of factors 1 and 2, on the other hand, involves us in a thorough examination of our basic concepts. If we consider classification, mazes, directions, California Mental Maturity Non-language, and possibly riddles (for in the latter, the usual cultural skills might be an obstacle to solution, by negative transfer) as most embodying sheer non-cultural relation eduction with little benefit from culture, then we are faced by the problem that two of these are strongly and one weakly shared with the second factor, which is tentatively identified as Visualization.

Nevertheless, our identification of Factor 1 as the fluid general ability, g_f, is aided by: (1) Its greater generality across other cognitive tests (loading, above $+\cdot 15$, no fewer than nine variables, compared to five for Factor 2); (2) Its substantial loading on intelligence performances that are not *only* perceptual in nature. At first glance, the interpretation of Factor 2 (Table 3) could be either that it is Reasoning, R, or Visualization, Vi (Thurstone, 1938; French, 1951; Horn, 1964).

TABLE 3

FACTOR PATTERN MATRIX FOR THE FOUR OF THE EIGHTEEN FACTORS STRONGLY INVOLVED IN COGNITIVE PERFORMANCES*

Nos	Variables Names	Factors			
		1 $g_{f(o)}$	2 R or V_1	3 V	4 N
	Cognitive				
1.	Substit. (Spd)	–06	03	$\boxed{44}$	07
2.	Substit. (low error)	00	–04	–00	09
3.	Classif.	$\boxed{41}$	$\boxed{67}$	02	–06
4.	Maze	$\boxed{20}$	$\boxed{23}$	08	13
5.	Sel. Obj.	$\boxed{36}$	06	–05	$\boxed{22}$
6.	Directions	$\boxed{52}$	–01	08	$\boxed{21}$
7.	Wrong Pic.	50	$\boxed{72}$	–01	07
8.	Riddles	$\boxed{20}$	03	03	$\boxed{25}$
9	Similar.	–16	11	05	$\boxed{49}$
10.	Cal MM (L)	$\boxed{45}$	–18	–06	00
11.	Cal MM (NL)	$\boxed{62}$	15	–02	–06
12.	Read. Word	06	13	$\boxed{48}$	07
13.	Read. Sent.	19	–12	$\boxed{66}$	00
14.	Read. Info.	$\boxed{24}$	–03	17	–02
15.	Read. Fol. Math.	–02	16	$\boxed{59}$	11
16.	Number Skill	08	–04	09	$\boxed{94}$
17.	Copying	07	–01	$\boxed{62}$	$\boxed{27}$
18.	Reading Ach.	12	–01	$\boxed{86}$	–08
19.	Number Ach.	03	05	–07	$\boxed{83}$
20	Draw (Intell.)	–07	19	–00	09
	Personality :†				
21	CP 1a	–11	–10	06	–02
22.	CP 1b	10	47	02	00
23.	CP 2a	–14	–08	04	–06
24.	CP 2b	14	12	07	–02
25.	CP 3a	03	02	11	–09
26	CP 3b	00	–04	–07	16
27.	CP 4a	–03	–17	54	02
28.	CP 4b	02	10	03	02
29.	CP 5a	–07	–11	–01	–15
30.	CP 5b	00	17	–04	–09
31.	CP 6a	07	–13	08	11
32.	CP 6b	50	–12	–05	–15
33.	CP 7a	10	–47	46	–02
34.	CP 7b	–00	16	36	–10
35.	CP 8a	–04	13	05	–34
36.	CP 8b	–03	–01	–04	–06
37.	CP 9a	06	–01	01	12
38.	CP 9b	–03	–07	–03	13
39.	CP 10a	–21	–06	01	01
49	CP 10b	–09	–24	–07	18
41.	CP 11a	–02	–09	04	–11
42.	CP 11b	–05	–09	01	05
43.	CP 12a	01	25	–01	–07
44	CP 12b	–18	01	–09	19
45.	CP 13a	–41	09	–08	10
46.	CP 13b	10	–01	–01	05

* Note that for economy of space only the four relevant factors, with cognitive ability loadings, are reproduced here. The remaining fourteen are clear doublets marking the thirteen known personality factors and an unknown fourteenth, also almost wholly in personality variance. Critical loadings for theory set in boxes.

† CP means Cattell-Peterson personality factor, from the questionnaire personality factor markers of their earlier study.

The issue is not unimportant, for according to present views, R is taken as a regular primary ability (possibly covering deductive and inductive reasoning ability), whereas Vi, by Horn's analysis (1964) is almost as broad as g_f and g_c, and, along with other possible sense area structuring abilities, could be on the same stratum, or slightly below, the *general* ability level If F2 is R, why does it not load more highly Following Directions, Riddles, and the traditional intelligence test items (Calif. M.M.) ? If, on the other hand, it is Vi, why does it not load more highly Perceiving Similarities and why is it so high on Classifications ? Only when the loading pattern of those factors has become more explored at 4–5 years can the decision be made. Because of the heavy involvement in Wrong Pictures, and some in Drawing, we lean a little to the Vi identification, but the arguments for R are probably equally good—and in the second order outcome better. Additionally, there is the possibility that through some insufficiency of items discriminating R and Vi (in hyperplane terms) at this age, F_2 is actually an unresolved *combination* of R and Vi.

Where then is the crystallized general ability factor ? In this rich setting of varied abilities, it is perhaps not surprising that it does not appear immediately at the first order. Instead, we find two clear-cut primary abilities ; first, in Factor 3, a very clear Verbal-reading ability factor (Variables twelve through fifteen and seventeen through eighteen, with slight overflow into substitution), and, second, in Factor 4, a very clear Number factor,with very high (·83, ·94) loadings on the two Number performances (but also slighter ones on the reasoning and exactitude types of performance, which might even suggest R).

The question now arises : " Will the g_c pattern appear at the next order, i.e., as a second order factor, loading all three of these primaries (or possibly omitting Vi, if such it is), but preserving its distinction from the fluid ability of Factor 1 ? " Parenthetically, the g_f factor is labelled $g_{f(o)}$ in Table 3, because if g_f is truly cognate with g_c we should expect first to recognize it formally, as such, in an appearance at the *second* stratum level. That is to say, we are adhering for the present to the defining of g_f and g_c as they have appeared before, at the *second* stratum (Cattell, 1963). As we shall argue later that the g_f factor picked up at the ' primary ' ability level is really a ' shorn ' fluid ability, lacking the loadings which it may have on the primaries, we have indexed it as $g_{f(o)}$—incomplete fluid ability pattern—making $g_{f(1)}$, later, the true span of fluid ability among all abilities. This stratum level issue is developed more fully in further discussing Diagram 1 (above) in Section 4 below.

Accordingly, in pursuit of the second stratum patterns, let us note that personality factor 18 in Table 3 which, by content, is a neurotic or introversive factor, has the interesting property of being at this age positively related to fluid ability and negatively to crystallized. This may mean that high natural ability in the young child increases susceptibility to conflict but that the educational forces which develop crystallized ability simultaneously work to reduce conflict. (Negative loadings in Table 3.) A second order analysis of the eighteen primary factors was begun with the factor correlations obtained from the very careful rotations made to simple structure at the first order. The Scree test, as Diagram 3 shows, indicated eight factors. Communalities were accordingly iterated to exactitude at three decimal places for eight factors and the resulting V_0 was rotated by Maxplane (Eber, 1966) to the first approach to simple structure. Thereafter, it was subjected to an extremely careful pursuit of blind simple structure by Rotoplot (Cattell and Foster, 1963) which eventually stablized unimprovably after twenty rotations at a ± 10 hyperplane count of 66 per cent. The R_f, V_0, and L (transformation) matrices are reserved at ADI under No. 222. The resulting factor pattern is shown in Table 4.

DIAGRAM 3

Decision on Number of Factor by Scree Test : Second Order.

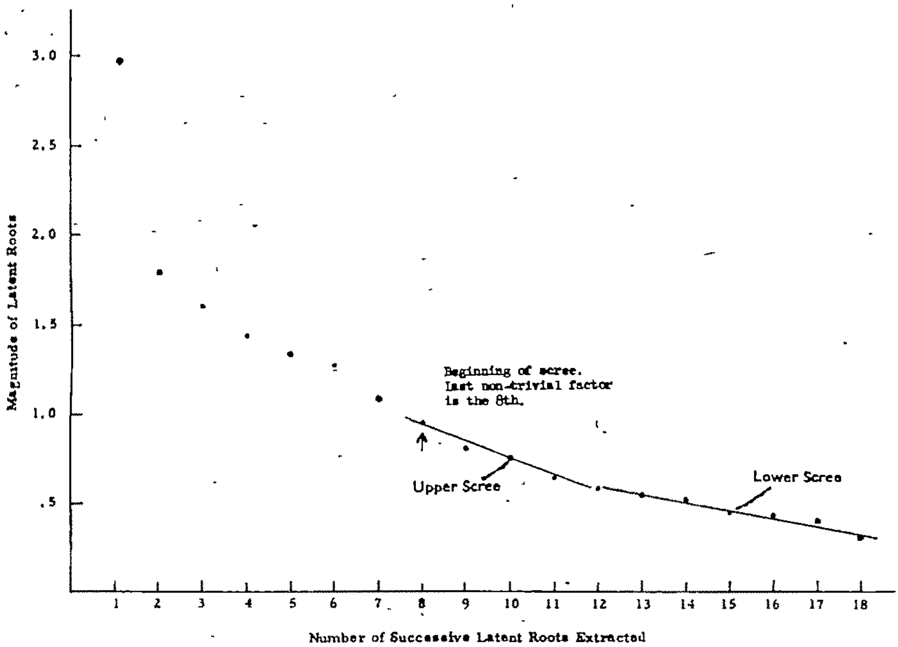


TABLE 4

Second Order Analysis : Factor Pattern.

	$g_{f(1)}$	$g_{c(1)}$	CPI	CPII	CPIII	CPIV	CPV	CPVI
Fluid Ability $g_{f(o)}$	$\boxed{58}$	—11	—03	—09	—09	14	01	—04
Vi. or Reason.	10	$\boxed{72}$	54	08	09	08	—05	02
Verbal (Read.)	—17	$\boxed{74}$	14	06	—05	—04	31	—53
Numerical	$\boxed{43}$	$\boxed{49}$	06	12	01	—06	05	04
CPF 1	14	—01	02	49	10	41	—08	09
CPF 2	04	—05	—08	—01	—10	—06	35	—01
CPF 3	07	—08	—05	03	10	14	—02	—65
CPF 4 (D)	—07	$\boxed{-38}$	—03	—02	—66	02	10	02
CPF 5 (I)	—13	—04	—10	06	—21	64	01	03
CPF 6 (E)	—12	$\boxed{-33}$	—05	01	—05	—07	—29	05
CPF 7 (O)	—00	—01	51	—01	05	01	—51	—07
CPF 8 (A)	08	$\boxed{50}$	—06	—02	—07	—08	—28	—31
CPF 9 (C)	—07	—09	—17	02	56	—15	18	—05
CPF 10 (H)	15	17	—06	—41	06	—04	—06	05
CPF 11 (M)	—03	$\boxed{-23}$	04	—33	41	05	—02	40
CPF 12 (Q_3)	01	02	11	—25	12	61	15	00
CPF 13 (Q_4)	—10	—02	—60	32	09	08	07	03
CPF (Bl. Dec.)	$\boxed{52}$	$\boxed{-53}$	09	—00	—01	—15	05	04

All decimal points have been omitted.

* CPF now refers not to Cattell-Peterson *markers*, but to the Cattell-Peterson first order personality factors *themselves*. The letters in parentheses after some of them indicate their very tentative 'content' identification (to be checked), with source traits in the 16 PF, HSPQ, CPQ, etc.

Now we see that the fluid general ability factor has retained its independence, though the loading (0·58) is not so high as to justify believing that $g_{f(o)}$ is the 'whole story' of the nature of $g_{f(1)}$. The second stratum $g_{f(1)}$ now appropriately loads N also (and, for some as yet unknown reason, a personality factor, possibly of anxiety engendered by the gift of high native ability). Meanwhile, the primaries V (verbal), N (numerical), and R-Vi (visual or reasoning), have come together in a single second order which has the meaning commonly assigned to Crystallized General Ability, g_c. At least, it clearly does so as far as V and N are concerned, through if Primary No. 2 is Visualization, not Reasoning, we would have to adjust other aspects of ability theory to a new phenomenon, namely, the inclusion of Visualization in Crystallized General Ability. Personality factors D. E. A. and H. load the latter as found elsewhere.

Beyond these two general ability factors lie six second order personality factors which, unfortunately, cannot be identified until two researches now in press (Cattell and Dreger, In press ; Cattell, Dreger, Coan and Baker, In preparation) are cross-checked. It is important to the present ability researches to have these identifications, because the fact that fluid ability positively affects personality factor CP3 and crystallized ability has a strong association with CP9, may or may not amplify and check our concepts of fluid and crystallized abilities. However, already we may note from Table 4 that whereas fluid general ability remains a relatively pure ability factor, crystallized intelligence begins to involve several personality factors (CPF 4, 6, 8 and 11, besides the last and unknown factor), such as one would expect if the realization of fluid ability in crystallized judgmental skills involved also temperament and interest influences. Parenthetically, the fact that $g_{f(1)}$ now shows some loading on N but not V may imply that V abilities have been growing for some years (since birth), whereas N is more recent, and generally rests only on a year's schooling, so that advance in it is more dependent on recent leads in $g_{f(1)}$.

Now, according to theory, and the previous findings with which it has been associated (Cattell, 1963 ; Horn, 1964, etc.) the next higher order analysis should show $g_{c(1)}$ to be the offspring of $g_{f(2)}$ by time and interest, i.e., a $g_{f(2)}$ corresponding to the individual's g_f level over the immediately preceding years should appear, loading g_c, but not itself loaded *by* g_c. That is to say, $g_{f(1)}$ and $g_{c(1)}$ in Table 4 should be positively correlated and on factoring should yield a factor heavily loading $g_{f(1)}$ but also accounting for an appreciable part of the variance in g_c. Since most of the V and N learning in these children is very recent, and the $g_{c(1)}$ is the product of $g_{f(1)}$ acting in very recent times, the g should heavily load the g_f of today.

To test this, the next higher order factoring was carried out. The scree (not again reproduced) gave four factors and the rotation yielded the factor pattern shown in Table 5.

The unmistakable result is that the fluid and crystallized second orders still fall on two ***different*** factors at this third stratum. However, they are no longer on their own, for crystallized ability has quite substantial associations with three second order personality factors—CPI, III, and VI—while fluid general ability has substantial association only with one, namely, CPVI. Before discussing the meaning of this we sought the evidence of a fourth order factor analysis. Here the scree test gave definite indication of three factors in four variables, which renders simple iteration to stable and unique communalities impossible. Accordingly, we employed a communality-estimating computer program by Nesselroade which employs a Jacobi eigenvector routine and iterates to obtain communalities convergent for the number of positive roots

TABLE 5

Third Order Analysis. Factor Pattern.

2nd Order Variable	Third Order Factors			
	$g_{f(2)}$	$g_{c(2)}$	CPα	CPβ
$g_{f(1)}$	$\boxed{85}$	–01	04	–08
$g_{c(1)}$	05	$\boxed{73}$	–04	01
CPF I*	00	$\boxed{-64}$	37	–66
CPF II	$\boxed{-33}$	–05	23	38
CPF III	03	$\boxed{-63}$	00	–41
CPF IV	–03	00	09	–54
CPF V	–01	–02	83	02
CPF VI	$\boxed{-80}$	$\boxed{68}$	00	–03

* Note.—As in the usual convention, Roman numerals are introduced to indicate that these personality factors are at the second stratum level.

TABLE 6

Fourth Order Analysis: Factor Pattern and Factor Correlations.
(a) Factor Pattern.

	$g_{f(3)}$	$CP\alpha_{(2)}$	$CP\beta_{(2)}$
$g_{f(2)}$	94	–06	13
$g_{c(2)}$	41	38	–12
$CP\alpha_{(2)}$	10	89	03
$CP\beta_{(1)}$	01'	–01	98

(b) Correlation Among the Fourth Order Factors.

	$g_{f(3)}$	$CP\alpha_{(2)}$	$CP\beta_{(2)}$
$g_{f(3)}$	1·00	·38	·02
$CP\alpha_{(2)}$	·38	1·00	—·12
$CP\beta_{(2)}$	·02	—·12	1·00

demanded. It *may* not be able to reach the given required number of positive roots, but the fact that it succeeds, as in this case, is additional support for that number being acceptable. The V_o, L, etc., are available at ADI under No. 333.

Regarding the rotational resolution in Table 6, and to some extent in Table 5, the reader must be reminded of the methodological point that as the number of variables diminishes, sampling problems become obtrusive. Roughly, one might say that when the number falls below, say, ten in a four factor space, the sample becomes insufficient to permit a specific simple structure position to be accepted with a high level of confidence, from a single investigation. Thus, we know that alternative positions exist to those in Tables 5 and 6 which are almost as good in one or two cases and conceivably could be the 'true' positions. Only additional experiments and, hopefully, some additional variables can raise the confidence level of the conclusion we are drawing from rotation in this last stratum. However, certainly the $g_{f(2)}$ and $g_{c(2)}$ of Table 5 correlate positively so that a single higher order embracing both is likely. And, in fact, the simple structure, unusually good at so high an order, does end with a g_f embracing g_c as general theory would lead us to expect (Table 6). It shows three 'variables' from the lower order each loading a higher order factor so highly (possibly unity, allowing for error of rotation) that CPα, CPβ and g_f substantially constitute the factors. Parenthetically, a two factor analysis of the four variables, iterating for communalities, yielded only a poor simple structure, suggesting that the Scree test is correct in negating a two factor resolution.

IV.—SUMMARY AND DISCUSSION.

(1) The theory of fluid and crystallized general ability factors has been briefly recapitulated as an introduction to this experiment, and some subhypotheses stated regarding differences of pattern to be expected at early and late ages, and with and without a large age span in the sample.

(2) The experiment measures 114, 5 to 7-year-old, boys and girls on twenty cognitive ability variables and twenty-four questionnaire, brief, paired, personality scales. The variables were strategically chosen in the cognitive field to separate fluid and crystallized ability (mainly by the Cattell Culture Fair Scale 1; the California Mental Maturity Test and Verbal and Numerical performances). The second (personality domain) set employed the Cattell-Peterson finding of thirteen personality factors at this age to provide a contrasting 'ground' of hyperplane stuff by which factors general to abilities could be accurately rotated, if necessary as far as the third and fourth orders.

(3) Using technical checks—notably (*a*) the Scree test for number of factors, (*b*) iteration of factoring to uniquely determined communalities, (*c*) unprejudiced (blind) rotation, initially by the Maxplane program, and, (*d*) at least fifteen to twenty further visual shifts (without knowledge of individual variables) by the Rotoplot program, to an unimprovable total simple structure count—eighteen factors were located at the first order, eight at the second, four at the third, and three at the fourth.

(4) Except for the clearly recognizable fluid ability factor, $g_{f(o)}$ (shorn of loadings on primaries, however) and the thirteen personality factors which appeared essentially as in the Cattell-Peterson study, the first order yielded no crystallized general ability, but only strong primary abilities—verbal, V, numerical, N, and a factor which could be either reasoning R or visualization, Vi, or a composite of both.

(5) At the second order, the ability primaries became organized in what is a typical crystallized general ability factor $g_{c(1)}$, which remains separate from a typical, extensive fluid ability factor, $g_{f(1)}$, now loading also a little of the numerical primary, and a neuroticism-introversion-like personality factor.

So far, this structure matches very closely what has been found in the only previous analyses (Cattell, 1963; Horn, 1964) meeting the above technical conditions. But at this point, one must discuss the phenomenon here encountered, and to be encountered again in a moment, in which a higher stratum factor (Cattell, 1965) seems apparently to make its appearance also at a lower order ($g_{f(1)}$ as $g_{f(0)}$ here). This can happen if other primaries which it should load are not present in the first order analysis except as single variables. One then has the alternative of 'stealing' some of the variance of these variables to appear as loadings on the second stratum factor, or of adding new variables to create the missing primaries and letting the second stratum factor properly show its loadings on these primaries at the next order. By this reasoning $g_{f(0)}$ has been described as a 'shorn' fluid ability factor, not having all its branches, and appearing at the same level as these branches.

The only circumstance here which does not fit this interpretation is the failure of $g_{f(0)}$ to have virtually a unit loading (actually it is $+0 \cdot 58$) on $g_{f(1)}$. However, $g_{f(1)}$, as Table 4 shows, loads only two out of eighteen other variables (and it *could* with slight rotation load only one). Consequently, the communality of $g_{f(0)}$, by reason of $g_{f(0)}$ standing almost alone, could be seriously underestimated. Thus, 0·58 is a biassed estimate of what might really reach unity. Parenthetically, one perceives in restrospect that a last small shift of $g_{f(0)}$ could have been made toward N which would have presented an alternative and perhaps equally good simple structure. The shorn $g_{f(0)}$ would then have been a little broader, but would have lost some loading on primary N at the next order. Fortunately, whichever way we choose to rotate it would not affect (except trivially) what we obtained at the next order.

(6) At the third order the structure is reduced to two general ability factors, indexed as $g_{f(2)}$ and $g_{c(2)}$ and two-third order personality factors, carrying the usual Greek letter indices, as CPα and CPβ, to indicate the third order. In the previous two studies, the two general ability factors have at this level come together as one 'historical' fluid ability factor, the crystallized ability factor $g_{c(2)}$ being revealed at the same time as a product of the historical fluid ability interacting with certain historical personality factors. ('Historical' will be explained in a moment.) The same resolution comes here, *but only at the fourth order*.

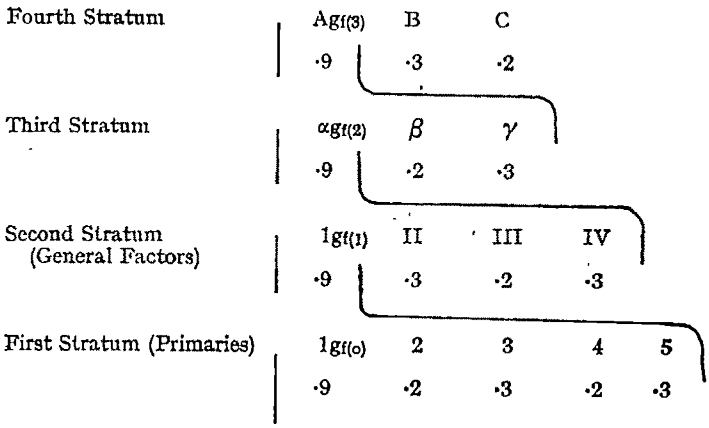
Two possibilities have to be considered to explain this: one merely mechanical, one involving an amendment to theory. The mechanical possibility is that we made a wrong estimate, by taking four factors instead of three, at the third order. It can only be repeated that the evidence by the usual standards showed four. We are in process, however, of recalculating with three only to see whether, as one would expect, this would yield a result not unlike our fourth order analysis here. It remains to be seen whether the general fit can be as good when one thus skips an order, and we doubt it.

To understand the alternative of fitting this best structural verdict by an amendment of theory, one needs to ask what meaning can be attached to the special mosaic (Cattell, 1965) we encounter in factor orders here. Stripped to its essentials it is shown in what we may call the hegemony mosaic or model in Diagram 4.

Here, as instanced by the contrast of a ·9 loading and loadings of ·2 and ·3 (a simpler picture of Tables 4, 5 and 6) a factor at each stratum accounts for virtually the whole of a factor at the next lower stratum. This degree of determination—i.e., of hegemony in determining a next lower stratum factor—justifies inferring some special conceptual continuity in the series of factors. In this case our first hypothesis is an historical continuity: $g_{f(3)}$ represents fluid ability as it operated and produced effects at some very early stage of the child's development; $g_{f(2)}$ represents it during the first year in nursery school, and $g_{f(1)}$ represents it as it stands functionally at the present moment. (An alternative for $g_{f(3)}$ is that it represents a single genetic influence which affects both fluid ability and certain temperament factors.)

DIAGRAM 4

THE HEGEMONY MOSAIC IN FACTOR STRATA.



Fourth Stratum	$Ag_{f(3)}$	B	C		
	·9	·3	·2		
Third Stratum	$\alpha g_{f(2)}$	β	γ		
	·9	·2	·3		
Second Stratum (General Factors)	$Ig_{f(1)}$	II	III	IV	
	·9	·3	·2	·3	
First Stratum (Primaries)	$1g_{f(o)}$	2	3	4	5
	·9	·2	·3	·2	·3

(NOTE—The individual α's, β's, etc. do not have same meaning as in other tables but Greek letters still stand for third stratum, and so on).

At each upward step in stratum, in any case, additional affected traits come into the pattern of the factor. But $g_{f(1)}$, the factor defined as the presently operating fluid ability, will be almost, though not entirely, restricted to primary abilities. By contrast, $g_{f(2)}$ has substantial loadings on one or two broad personality factors, because the intelligence of the early years interacted with environment to develop certain personality factors too.

(7) The evidence of Table 6, at the fourth order, finally yields the same conclusion as from the pioneer researches: crystallized general ability is revealed as the product of investment of the earlier fluid ability, $g_{f(3)}$: of some personality traits, presumably of persistence and application, labelled CPA, and of time. But (unless the above mechanical, statistical estimate failure explains all) why has this relationship come at a stratum higher than usual?

Let it be noted that the same hegemony model and lineal descent (in $g_{c(1)}$ and $g_{c(2)}$) holds, in this experiment, for the g_c's as for the g_f's. The crystallized ability in the child today, at 5 or 6 years, $g_{c(1)}$ (which, incidentally includes a pattern of personality and social skills—CP4, CP8 and 14—as well as cognitive skills) is the product and descendant of the $g_{c(2)}$ pattern of the previous few years. This tendency of crystallized ability to beget crystallized ability through 'aids' has been expressly part of the fluid and crystallized general ability theory (Cattell, 1963, 1965). But why should the old g_c's influence on the new g_c be greater here than in more mature years? Presumably because the individual differences in opportunity to pick up aids are greater and more accidental over two or three extremely important early developmental years than when steadily accumulated over many years in a standardized school environment. However, even on this theory, of more active influence of $g_{c(2)}$ in generating individual differences in $g_{c(1)}$ in early childhood than later, we should expect $g_{f(2)}$ to have contributed something more than the ·05 value in Table 5, to the $g_{c(1)}$ score during the immediately succeeding period.

Future experiment should aim at accurate determination of this value. It should obviously also check the general theory here by more than the cross sectional approach necessary in a first attack. A measuring and remeasuring of the same children over a one-year and two-year interval offers a means of checking the development sequences here inferred.

Recently, there has been some tendency to try to re-orient the concepts of some British psychologists (outstandingly of Vernon, 1950)—which contrast practical ability, k, with educational verbal ability, v : ed—to fit the findings appearing around the fluid and crystallized ability concepts. But this attempt to creep from one conceptual position to the other, without clear rejection of false associations and different technical assumptions is merely asking for confusion. The resemblance of the two dualities is only superficial and gross. Fluid and crytallized abilities are oblique factors : k and v : ed are orthogonal. The g_f and g_c theory supersedes and abolishes the need for the Spearman g altogether, whereas the k and v : ed theory retains g. The crystallized ability factor, g_c, covers alike mechanical, practical, numerical and verbal skills, whereas the k and v : ed divide them into two different categories. There are, besides, a whole set of connotations of the fluid and crystallized ability theory in physiological, cultural and developmental fields which do not belong to the theory of the k and v : ed dichotomy.

ACKNOWLEDGEMENTS.—The writer wishes to express his great indebtedness to research assistants John Nesselroade and Larry Killian who gathered the measurement data, to the principals of the schools which gave their facilities, and to Jack Ford who carefully performed the prolonged computer analyses. This investigation was supported in part by Public Health Service Research Grant No. 1733-10 from the National Institute of Mental Health.

V.—REFERENCES.

BURT, C. L. (1941). *Factors of the Mind.* London : Univ. of London Press.

CATTELL, R. B (1941). Some theoretical issues in adult intelligence testing. *Psychol. Bull*, 38, 592.

CATTELL, R. B. (1950) *Culture Fair Intelligence Test Scale* 1. Champaign, Ill. : Instit. Pers. and Abil. Test.

CATTELL, R. B. (1965). *The Scientific Analysis of Personality.* Baltimore : Penguin Books.

CATTELL, R. B. (1963). Theory of fluid and crystallized intelligence : a critical experiment. *J. Educ. Psycol.*, 54, 1-22.

CATTELL, R. B. (1965). Higher order factor structures and reticular-vs-hierarchical formulae for their interpretation In C Banks and P. L. Broadhurst (Ed.), *Stephanos : Studies in Psychology Presented to Cyril Burt.* London : Univ London Press, Chap. 14, pp. 223-266.

CATTELL, R. B. (Ed.) (1966). *Handbook of Multivariate Experimental Psychology.* Chicago, Ill. : Rand McNally.

CATTELL, R B., and DREGER, R. M. (1967). *Personality Structure in 4-5-year-olds in the Q-data Medium.* (In press.)

CATTELL, R. B., DREGER, R. M , COAN, R. B., and BAKER, R. (1967). Identification, by ESPQ relations, of checked personality source traits (Q-data) in 4-5-year-olds (In preparation).

CATTELL, R. B., and FOSTER, M. J. (1963). The rotoplot program for multiple, single plane, visually -guided rotation. *Behav. Sci.*, 8, 156-165.

CATTELL, R. B., and PETERSON, D. K. (1958). Personality source traits factored from the questionnaire responses of 4-5-year-olds. *Adv. Publ. No.* 15, Lab. of Pers. Anal. and Group Behav., Univ of Ill., Urbana, Ill.

EBER, H. W. (1966). Toward oblique simple structure . maxplane *Multiv. Behav. Res.*, 1, 112-125.

French, J W. (1951). The description of aptitude and achievement factors in terms of rotated factors. *Psychol. Monogr.*, No. 5.

Hebb, D. O. (1942). The effect of early and late brain injury upon test scores, and the nature of normal adult intelligence. *Proc. Amer. Phil. Soc.*, 85, 275-292.

Horn, J. L. (1964). *Fluid and Crystallized Intelligence: A Factor Analytic Study of the Structure Among Primary Mental Abilities.* Unpublished Doctoral Thesis. Urbana Ill., Univ. of Illinois Library.

Horn, J. L., and Cattell, R. B (1966) Age differences in primary mental ability factors *J. Gerontology*, 21, 2, 210-220.

Spearman, C. (1904). General intelligence, objectively determined and measured. *Amer. J. Psychol*, 201-293.

Thurstone, L. L. (1938). *Primary Mental Abilities.* Chicago · Univ of Chic. Press.

Vernon, P. E. (1950). *The Structure of Human Abilities.* London: Methuen.

THE SELF-CONCEPT IN THE PSYCHOLOGICAL GUIDANCE OF ADOLESCENTS

BY HAMED A. S. ZAHRAN
(*Faculty of Education, Ain Shams University, Cairo*)

SUMMARY. An attempt was made to relate the Self-Concept to the whole personality structure and to indicate the significance of the relationship for psychological guidance of adolescents. A large number of Self-Concept, counselling and adjustment, intellectual, personality, perceptual, social and environmental tests covering 130 variables were given to a group of 170 adolescents.

Cluster analysis (which was confirmed by factor analysis) revealed a main cluster of five variables on the basis of which the sample was divided into three, a Positive Self-Concept group, a Negative Self-Concept group and a middle group which was labelled ' Normal Control.' Analysis showed these three groups to be differentiated by a large number of variables. There were also sex differences and differences between extreme groups within each sex

I.—INTRODUCTION.

THE Self-Concept may be defined as: an organised, learned, cognitive and unitary configuration of conscious perceptions, conceptions and evaluations by the individual, of his Self as he actually is (Perceived-Self), as others are supposed to see him (Other-Self), and as he would most like to be (Ideal-Self). The function of the Self-Concept is motivating, integrating, and organising the world of experience at the centre of which the individual exists and thus regulates behaviour. It develops as a product of social interaction with the inner drive to Self-maintainance, and though it is fairly consistent, yet it is changeable under certain conditions. The most important contemporary theory of Self is that of Carl Rogers (1951-1961) of which the Self-Concept constitutes the core. Its main characteristics are that it pictures the end-point of personality development, as being a basic congruence between the phenomenal field and experience and the conceptual structure of the Self—a situation which, if achieved, would represent freedom from internal conflict and anxiety. The main importance and relevance of Rogers's theory is that it developed from his experience of counselling and pyschotherapy and that it deals systematically with the main characteristics of the fully functioning human organism who may, under certain circumstances, become relatively psychologically maladjusted, and under certain conditions within the process of client-centred therapy can become again a fully functioning person.

The Self-Concept is thus of central importance in psychological guidance, as distinct from educational and vocational guidance. It is the responsibility of psychological research to provide the counsellor or the therapist with the main features of the Self-Concept and the significant characteristics of positive and negative Self-Concepts. This is what was attempted in the present research.

Many studies have already been carried out into the Self-Concept, but most of these have used small samples with different characteristics and have covered only a small number of variables and with different techniques. Moreover, contradictory views and findings were often obtained. Other researches have provided no systematic or generalised results which would be helpful to counsellors and therapists by pointing out the psychological characteristics of individuals holding positive or negative Self-Concepts as compared with normal controls. All these remarks point to the need for replication and for grouping together larger numbers of variables in one larger sample of adolescents.

The emergence of this problem together with the findings of previous research led the writer to formulate a series of main and secondary hypotheses to be tested in his experiment. These concerned the psychological characteristics of individuals holding positive Self-Concepts, individuals holding negative Self-Concepts, and normal or average individuals, and the relationships of the Self-Concept to other personality variables.

II.—THE EXPERIMENT.

(*a*) *The Subjects.* A large battery of tests was applied to a sample of 220 adolescent boys and girls (13–15 years of age), in a London secondary modern school. Complete results were obtained from eighty girls and ninety boys. Most of the children came from working class homes.

(*b*) *The Variables.* A number of Self-Concept, Counselling and Adjustment, Intellectual, Personality, Perceptual, Social, Environmental, and other tests were employed to yield 130 variables.

These tests were*:

1. Abstraction Test of Intelligence (Cf. Vernon and Parry, 1949).

Selective Vocabulary Test (Constructed by Slater, 1944):

2. Vocabulary.
66. Masculinity/Femininity, i.e., Identification with same *vs.* opposite sex.

Minnesota Counseling Inventory (Constructed by Berdie and Layton, 1957):

3. Defensiveness and display of Social Desirability.
4. Healthy family relationships *vs.* Maladjusted family relationships.
5. Social maturity *vs.* Social maladjustment.
6. Emotional stability *vs.* Emotional instability.
7. Reliability, responsibility, conformity *vs.* Irreliability, irresponsibility, rebelliousness.
8. Realism, establishing sound relationships with groups *vs.* Secretiveness, withdrawal, sensitivity and difficulty in making friends
9. Good morale, cheerfulness *vs.* Poor morale, depression.
10. Leadership *vs.* Ineptness in social relations, avoidance of participation in groups.
11. Personality adjustment *vs.* Personality maladjustment.

Sentence Completion Test (Constructed by the writer, Zahran, 1964):

12. School adjustment *vs.* School maladjustment.
13. Social adjustment *vs.* Social maladjustment.
14. Home adjustment *vs.* Home maladjustment.
15. Emotional adjustment *vs.* Emotional maladjustment.
16. Physical health adjustment *vs.* Physical health maladjustment.
17. Overall personal adjustment *vs.* Overall personal maladjustment.

Self Rating Q-Sort (Constructed by the writer):

18. Self-Ideal congruence.
19. Physical fitness.

---

* For the detailed description of the tests, see: Zahran, H. A. S. (1966).

20. Academic capability.
21. Personal fitness.
22. Social adaptation.
23. Humanitarian social values.
24. Rich interests.
25. Variety of interests.
26. Self-confidence.
27. Satisfaction with present status *vs.* Eagerness for attainment of status.
28. Sensitivity to concerns of others *vs.* Egocentrism.
29. Healthy family relationships.
30. Self-Concept, general positive indications.†
31. Sophistication-of-Body-Concept Drawings (Adapted by the writer from Witkin, *et al.*, 1962).
32. Spatial Test 1 (Constructed by Macfarlane Smith, 1959).
33. Embedded-Figures Test of Field Independence (Adapted by the writer from Gottschaldt, 1926, and Witkin, 1950. See also Witkin, *et al.*, 1962, and Vernon, 1965).

Jr.-Sr. High School Personality Questionnaire (Constructed by Cattell and Beloff, 1962).

34. Cyclothymia, sociability, warmth.
35. General intelligence.
36. Ego strength, emotional stability.
37. Security *vs.* Excitability.
38. Submissiveness *vs.* Dominance.
39. Surgency, cheerfulness.
40. Super-ego strength.
41. Parmia, gregariousness, boldness.
42. Harria, hardness, realism, self-sufficiency, *vs.* Premsia, dependence, femininity
43. Zeppia, liking group action, vigorousness *vs.* Coasthenia, individualism, neurasthenia.
44. Confident adequacy, self-security *vs.* Guilt proneness, insecurity.
45. Self-sufficiency.
46. Strength of self sentiment.
47. Low ergic tension, relaxed *vs.* High ergic tension, tense, excitable, frustrated.
48. Extraversion.
49. Mental health, good adjustment *vs.* Anxiety.
50. Authoritarianism Scale (Adapted by the writer from Adorno, *et al.*, 1950).
51. Prejudiced-Stereotyped Accusation Scale (Constructed by R. Pitts).

Sociometric Test (Constructed by the writer):

52. High sociometric choice.
53. Low sociometric rejection.
54. High leadership choice.

---

† Sets of statements in these various areas were sorted by the children for Perceived-Self and then for Ideal-Self The correlation between the two sortings represented the Self-Ideal congruence index score for each subject.

55. Low Leadership rejection.
56. High sociometric popularity status.
57. Articulateness of Experience Regarding Everyday Environment[*] (Adapted by the writer from Witkin, *et al.*, 1962).

Factorial Interest Form (Constructed by Sandall, 1961) :

58. Rural-practical interest.
59. Social-display interest.
60. Humanitarian interst.
61. Entertainment interest.
62. Physical interest.
63. Literature interest.
64. Aesthetic interest.
65. Scientific interest.

Semantic Differential (Constructed by the writer. See also Osgood, *et al.*, 1957) :
67-126. Four factorial labels.[*]

(Valuable, Active, Potent, and Calm *vs.* Excitable), applied to fifteen concepts (Myself now, My father, School, Social activities, England, My mother, Teachers, Myself in ten years time, Boys, My dreams, Family life, The world we live in, My future job, Girls, and Religion).

Academic Achievement (Taken from school records) :

127. English.
128. Mathematics.
129. Science.

Physical Somatometry (Following procedures by Parnell, 1958, and Tanner, 1964) :
130. Phsyical somatotype.[†]

(*c*) *Pilot Study*. The tests constructed or adapted by the writer were subjected to a pilot run on a sample of ninety-three secondary modern school pupils, aged 13–15, before they were given to the experimental sample in final form. These tests were the Self Rating Q-Sort, the Authoritarianism Scale, Articulateness of Experience Regarding Everyday Environment, and the Embedded Figures Test. This pilot study was useful in clarifying the feasibility, the potential adequacy of testing procedure, and in refining and reformulation of some test items and presentation procedures.

All the tests were administered as group tests by the writer.

III.—ANALYSIS OF DATA.

Having scored all the tests,[‡] the first step was to factor analyse the data obtained from the Semantic Differential for each sex separately. In both analyses, four significant factors were extracted and rotated into simple structure. These were labelled :

---

[*] The purpose of such an assessment was to obtain information about children's structuring and articulation of experience in 'real-life.' High scorers tend to experience the world and themselves in a relatively clear and structured fashion. Low scorers, on the other hand, tend to experience the world and themselves in a relatively unclear and unstructured fashion

[*] As extracted from the factorial analysis of the Semantic Differential (Zahran, 1965).

[†] Data from Physical Somatometry were incomplete, hence there was no clear or psychologically feasible trend.

[‡] Conversion tables were used in scoring Abstraction Test of Intelligence, Spatial Test 1, and percentile scores were calculated for the Selective Vocabulary Test (M/F).

I. General Evaluative factor.
II. Activity factor
III. Potency factor, and
IV. Excitability factor . (See Tables 1 and 2).

The factorial structures for both sexes were found to be essentially similar. Factor scores were then calculated along the four factors and entered into the main score matrix containing 129 variables and 170 subjects, eighty girls, and ninety boys.

TABLE 1 .

ROTATED PRINCIPAL COMPONENTS LOADINGS (SEMANTIC DIFFERENTIAL, GIRLS).

Scales	Loadings for 4 Components				
	I	II	III	IV	h^2
1 Good-bad	0·64	—0·10	0·29	—0·10	0·51
2. Active-passive	0·25	0·80	—0·14	0·06	0·73
3. Big-little	0·24	0·13	0·61	0·14	0·47
4 Happy-unhappy	0·63	0·25	—0·17	0·21	0·53
5. Quick-slow	0·54	0·47	0·14	—0·06	0·54
6. Wise-silly	0·68	0·08	0·00	—0·37	0·61
7. Fair-unfair	0·75	0·04	—0·06	0·00	0·56
8. Excitable-calm	0·03	0·29	—0·07	0·82	0·76
9. Masculine-feminine	—0·16	—0·05	0·74	—0·07	0·58
10 Kind-cruel	0·75	—0·08	—0·24	—0·08	0·63
11. Tense-relaxed	—0·59	0·34	0·21	—0·16	0·52
12. Strong-weak	0·53	0·15	0·44	—0·06	0·50
13 Interesting-dull	0·72	0·03	0·16	0·21	0·59
14. Honest-dishonest	0·71	—0·04	—0·20	—0·19	0·58
15. Successful-unsuccessful	0·70	0·14	—0·04	—0·18	0·54

TABLE 2

ROTATED PRINCIPAL COMPONENTS LOADINGS (SEMANTIC DIFFERENTIAL, BOYS).

Scales	Loadings for 4 Components				
	I	II	III	IV	h^2
1. Good-bad	0·70	0·25	—0·13	0·05	0·57
2. Active-passive	0·38	0·65	—0·01	0·01	0·57
3. Big-little	0·35	0·45	0·36	—0·14	0·47
4. Happy-unhappy	0·73	0·24	—0·03	0·07	0·60
5. Quick-slow	0·53	0·54	0·19	—0·01	0·60
6. Wise-silly	0·74	0·19	0·06	—0·12	0·60
7 Fair-unfair	0·79	0·15	0·06	—0·03	0·65
8. Excitable-calm	—0·02	0·22	0·02	0·95	0·95
9 Masculine-feminine	0·25	0·03	0·86	0·02	0·80
10 Kind-cruel	0·82	—0·08	—0·15	—0·01	0·70
11. Tense-relaxed	—0·59	0·43	—0·10	—0·16	0·57
12 Strong-weak	0·65	0·18	0·47	—0·05	0·68
13. Interesting-dull	0·68	0·24	0·20	0·14	0·58
14 Honest-dishonest	0·75	—0·15	—0·06	—0·11	0·60
15. Successful-unsuccessful	0·72	0·13	0·16	—0·02	0·56

The score matrix was then subjected to a series of statistical analyses to calculate means, standard deviations, and correlation coefficients,* for the whole sample.

From the matrix of correlation coefficients, cluster analysis revealed a main cluster called ' Self-Concept,' which included five variables :

18. Self-Ideal congruence (from Q-Sort).
46. Strength of self sentiment (from High School Personality Questionnaire).
11. Overall personal adjustment (from Sentence Completion Test).
4. Healthy family relationships (from Minnesota Counseling Inventory).
7. Conformity, reliability, responsibility (from Minnesota Counseling Inventory).

The last two variables were combined together as they were derived from one test ' MCI.'

The Self-Concept cluster was then confirmed by factor analysis of a reduced score matric containing only seventy-seven variables after eliminating some dependent variables and some variables which did not contribute much to the main Self-Concept cluster. In this factor analysis, six interpretable principal components were extracted. These were labelled :

I. Self-Concept factor,
II. Interests factor,
III. Ability factor,
IV. Sociometric Status factor,
V. Social factor, and
VI. Popularity factor (Cf. Zahran, 1966).

A Self-Concept Index was established by combining the normalised standard scores of the four Self-Concept variables.

On the basis of the Self-Concept Index scores, the sample was broken down into three main groups for each sex separately. The top third of the sample scoring the highest on the Self-Concept Index was labelled the ' Positive Self-Concept ' group. The bottom third of the sample scoring the lowest on the Self-Concept Index was labelled the ' Negative Self-Concept ' group. The middle group comprising the rest of the sample scoring moderately on the Self-Concept Index and falling between the two extremes, was labelled the ' Normal Control ' group.

Analysis of variance was applied for testing the significance of differences between these groups,† and correlations were also calculated between the Self-Concept Index scores and all other variables and these closely confirmed the evidence from the *t* tests.

IV.—RESULTS.

The most strongly differentiating characteristics, which gave significant differences between each extreme group and the middle group were :

Form Minnesota Counseling Inventory :

6. Emotional stability.
8. Realism, establishing sound relationships with groups *vs.* Secretiveness, withdrawal, sensitivity and difficulty in making friends
11. Personality adjustment *vs.* Personality maladjustment.

---

* Cf. Zahran (1966).

† For details, see Zahran (1966).

From Sentence Completion Test :

15. Emotional adjustment.

From Q-Sort :

27. Satisfaction with present status *vs.* Eagerness for attainment of status

From High School Personality Questionnaire :

40 Super-ego strength *vs.* Dependent character.
49. Mental health, good adjustment *vs.* Anxiety.

Subjects holding PSCs scored the highest (the positive pole) while those holding NSCs scored the lowest (the negative pole), and the NCs normally scored moderately on each trait.

Taking the Self-Concept Index as a valid measure of acceptance of Self, it was found to be positively and significantly correlated with other variables indicating acceptance of others, on the one hand, and acceptance by others on the other. These findings reinforce the contention that Self-acceptance and understanding is a principal dimension in the process of personal adjustment and readjustment.

These results support Rogers's contention that positive and accepting Self-Concepts are associated with healthy psychological adjustment. Consequently, distortions in the Self-Concept may be safely taken as signs and symptoms of psychological maladjustment. Scores on these variables should be of considerable use to the psychological counsellor.

Significant *F*-ratios between the two extreme groups were obtained on the following additional characteristics :

9. Good morale, cheefulness *vs.* Poor morale, depression.
12. School adjustment.
13. Social adjustment.
14. Home adjustment.
16. Physical health adjustment.
20. Academic capability.
23. Humanitarian social values.
28. Sensitivity to concerns of others *vs.* Egocentrism.
29. Healthy family relationships.
36. Ego strength, emotional stability *vs.* Ego weakness, neuroticism.
37. Security *vs.* Excitability
38. Submissivness *vs.* Dominance.
44. Confident adequacy, self-security *vs.* Guilt proneness, insecurity.
47. Low ergic tension, relaxed *vs.* High ergic tension, tense, excitable, frustrated.
51. Non-prejudiced non-stereotyped thinking (Boys sub-groups).
57. Cognitive clarity.
63. Literature interest (Boys sub-groups).
72. Father, high *vs.* low evaluation.
75. School, high *vs.* low evaluation.
83. England, high *vs.* low evaluation.
91. Teachers, high *vs.* low evaluation.
95. Self when grown up, high *vs.* low evaluation.

96. Self when grown up, high *vs.* low potency.
106. Dreams, calm *vs* excitable.
111. The world, high *vs.* low evaluation.
119. Girls, high *vs.* low evaluation.
121. Girls, high *vs.* low potency.
123. Religion, high *vs.* low evaluation.
125 Religion, high *vs.* low potency.
129. Good academic achievmeent ' Science.'

Subjects holding PSCs scored the highest while those holding NSCs scored the lowest on the above characteristics.

In the following additional characteristics the PSC group scored significantly higher than the NC group :

20. Academic capability.
28. Sensitivity to concerns of others *vs.* Egocentrism.
29. Healthy family relationships
38. Submissiveness *vs.* Dominance.
47. Low ergic tension, relaxed *vs.* High ergic tension, tense, excitable, frustrated.
57. Cognitive clarity.
82. Social activities, calm *vs.* excitable.

Other characteristics similarly differentiated significantly between the NC and NSC groups :

12. School adjustment.
13. Social adjustment.
14. Home adjustment.
36. Ego strength, emotional stability *vs.* Ego weakness, neuroticism.
37. Security *vs.* Excitability.
72. Father, active *vs.* passive.
83. England, high *vs.* low evaluation.
91. Teachers, high *vs.* low evaluation.
95. Self when grown up, high *vs.* low evaluation.
111. The world, high *vs.* low evaluation.
119. Girls, high *vs.* low evaluation.

These psychological characteristics which differentiated the PSC and the NC groups were different from those which differentiated the latter and the NSC group, showing that the two extremes were in some respects qualitatively different from each other.

The rest of the variables failed to differentiate the three groups or even any pairs. None of them was significantly correlated with the Self-Concept Index. Among these variables were some which showed significant sex differences (which accounted for most of the within group variance).

Significant sex differences were also found on some variables which differentiated between two of the groups and were at the same time positively and significantly correlated with the Self-Concept Index. Thus, girls scored higher than the boys on :

23. Humanitarian social values.
28. Sensitivity to concerns of others *vs.* Egocentrism.
57 Cognitive clarity.
91 Teachers, high *vs* low evaluation.
119. Girls, high *vs.* low evaluation.
121. Girls, high *vs.* low potency
123. Religion, high *vs.* low evaluation.
125. Religion, high *vs.* low potency.
129. Good academic achievement ' Science.'

The boys, on the other hand, scored higher than the girls on :

36. Ego strength, emotional stability *vs.* Ego weakness, neuroticism.
44. Confident adequacy, self-security *vs.* Guilt proneness, insecurity.
47. Low ergic tension, relaxed *vs.* High ergic tension, tense, excitable ,frustrated.
49. Mental health, good adjustment *vs.* Anxiety.
106 Dreams, calm *vs.* excitable.

The correlations between the Self-Concept Index and the other variables showed that, taking the total sample into account, the Self-Concept Index was positively and significantly correlated with the following variables :

3. Defensiveness and display of Social Desirability.
5. Social maturity.
6. Emotional stability.
8. Realism, establishing sound relationships with groups.
9. Good morale, cheerfulness.
10. Leadership.
11. Personality adjustment
12. School adjustment.
13. Social adjustment.
14. Home adjustment.
15. Emotional adjustment.
16. Physical health adjustment.
20. Academic capability.
22. Social adaptation (NC, NCG, NSCB)
23. Humanitarian social values.
25. Limitedness of interests.
27. Satisfaction with present status.
28. Sensitivity to concerns of others.
29. Healthy family relationships.
36. Ego strength, emotional stability.
37. Security.
38. Submissiveness.

39. Desurgency, depression.
40. Super-ego strength.
42. Premsia, dependence, femininity (PSCG, NC).
44. Confident adequacy, self-security.
47. Low ergic tension, relaxed.
49. Mental health, good adjustment.
51. Non-prejudiced, non-stereotyped thinking.
57. Cognitive clarity.
72. Father, active.
75. School, valuable.
79. Social activities, valuable (PSC, PSCG, NC, NCG).
83. England, valuable.
87. Mother, valuable.
91. Teachers, valuable.
95. Self when grown up, valuable.
106. Dreams, calm.
107. Family life, valuable.
111. The world, valuable.
119. Girls, valuable.
121. Girls, potent.
123. Religion, valuable.
125. Religion, potent.
129. Academic achievement ' Science.'

All these correlations were in the predicted direction. The majority of the variables correlated with the Self-Concept were also differentiating variables among the groups. Noticeable among them is the constant relationship between the Self-Concept and adjustment. The counsellor or therapist adopting the client-centred method may place more reliance on these variables which were significantly correlated as well as differentiating variables.

Some variables were significantly correlated with the Self-Concept Index in one group and not the other two. Special attention may be drawn to these which discriminated within the NSC group only :

24. Lack of interests.
31. Primitiveness of body concept.
113. The world, impotent.
116. Future job, active.

Of these, Primitiveness of Body Concept Drawings is a non-phenomenal (projective) test. It also gave a significant negative correlation with Self-Ideal congruence (No. 18). This suggests that for a complete understanding of the Self-Concept account should be taken of conscious and unconscious aspects particularly when the results show that we are dealing with a NSC.

The Self-Rating Q-Sort not only provided the Self-Ideal Congruence Index (No. 18), but also scores or Self-ratings on various traits (19-30). The correlations of some of these Self-ratings with other variables indicate that this is a particulary useful instrument in counselling.

No. 20—Academic Capability correlated positively and significantly with :

12. School adjustment.
24. Rich interests.
63. Literature interest.
75. School, valuable.
77. School, potent.
91. Teachers, valuable.
127. Academic achievement ' English.'
128. Academic achievement ' Mathematics.'

No. 23—Humanitarian Social Values correlated with :

27. Satisfaction with present status.
28. Sensitivity to concerns of others.
60. Humanitarian interest.

No. 27—Satisfaction with present status correlated with :

4-11. All sections of the Minnesota Counseling Inventory.
12. School adjustment.
14. Home adjustment
15. Emotional adjustment.
17. Personal adjustment
18. Self-Ideal congruence.
20. Academic capability.
21. Personal fitness.
23. Humanitarian social values
28. Sensitivity to concerns of others.
29. Healthy family relationships.
30. Self-Concept, general positive indications.
33. Field independence.
37. Security.
40. Super-ego, strength.
43 Zeppia, liking group action, vigorousness.
46. Strength of self-sentiment.
49. Mental health, good adjustment.
75, 79, 83, 87, 91, 107, 111, 123. High evaluation of School, Social activities, England, Mother, Teachers, Family life, The World, and Religion.
125. Religion, potent.
129 Academic achievement ' Science.'

No. 28—Sensitivity to Concerns of Others correlated with :

7. Reliability, responsibility, conformity.
8. Realism, establishing sound relationships with groups.
13. Social adjustment.
23. Humanitarian social values.

38. Submissiveness.
51. Non-prejudiced, non-stereotyped thinking.
60. Humanitarian interest.
99. Boys, valuable.
112. The world, active.

No.29—Healthy Family Relationships correlated with:

4. Healthy family relationships (Minnesota Counseling Inventory).
14. Home adjustment.
57. Cognitive clarity.
71. Father, valuable.
87. Mother, valuable.
107. Family life, valuable.

On the basis of these results, the present study supports the methodological usefulness of the self-rating technique, but recommends the use of some other techniques together with it.

The results also showed that the use of the Self-Concept Index as a basis for breaking the sample down into the three groups and their subgroups, was more adequate than just depending on one single test (Self-Ideal congruence). This was confirmed by the relatively small number of variables which correlated with Self-Ideal congruence compared with the larger number of variables correlated with the Self-Concept Index in the predicted direction. This was also confirmed in the cluster analysis and the second factorial analysis. Moreover, the factorial analysis of the Semantic Differential confirmed the multi-dimensionality of the Self-Concept as it was found that each of the four extracted factors contributed to the Self-Concept. It is important then that the multi-dimensionality of the Self-Concept should be recognised in Self-Concept studies.

It is sometimes argued that self-ratings and answers to personality inventories measure little but Social Desirability. This possibility was studied by observing the correlations for No. 3—Defensiveness and display of Social Desirability in the Minnesota Counseling Inventory. Although this score did correlate at the 0.05 level with the Self-Concept Index it did not differentiate among the three groups or even between the two extreme groups. Thus distorting effect of Social Desirability was not found to be as great as it is often emphasised in self-rating techniques. It seems likely, however, that the greater the agreement between a child's self-ratings and other more objective descriptions of him, the less defensive he will be and the more adequate will be his personal adjustment and *vice versa*. Probably it is only among the extreme NSC holders that distortions due to Social Desirability need to worry us both in testing and clinical practice. The Social Desirability tendency is not unique to the measurement of the Self. In personality testing we find other distorting factors among which are denial, faking, halo, etc.; all should be allowed for both in research and practice. It is thus desirable, for safety's sake, to include a Social Desirability scale, together with self-rating tests in Self-Concept studies.

In the Semantic Differential, the fifteen concepts ranged in terms of evaluation from top to bottom as follows:

(*a*) Family concepts,
(*b*) Self concepts,
(*c*) Social concepts,

(*d*) Environmental concepts, and
(*e*) School concepts.

This seems a logical order of the concepts. Notice the sense of belonging to the family in an inseparable way, and the close relationship between Self evaluation and the evaluation of family relationships and home adjustment. Religion is the least evaluated concept for all groups. None of the concepts was perceived negatively in terms of evaluation. Taking the total sample into account, only Religion, which was the least evaluated concept, was perceived as slightly passive. On the potency factor, Mother, and Girls were perceived as slightly impotent. All the other concepts were perceived as ranging from slightly to highly valuable, active, and potent. These results point not only to the importance of Self, but also to the extreme importance of the home environment and the social environment. The implications of these findings are that environmental treatment and manipulation should go hand in hand with counselling, starting from the immediate environment of the home and then outward to the wider society, etc.

Finally, factor analysis II classified the main variables in the whole battery of tests into six principal components. The Self-Concept factor included not only Self-Concept variables, but also some counselling, adjustment and personality variables. The social factor, together with the sociometric status and popularity factors, included social as well as environmental variables. The interests factor included different interest areas. Positive Self-Concept and personality adjustment proved to be highly associated, the presence of one may be taken as significant evidence for the other. It is noticeable that interests, which are commonly very important in vocational guidance, failed to gain a comparable status in psychological guidance.

V.—CONCLUSIONS.

A consistent and valid index of positive *vs.* negative Self-Concept can be derived from Q-sort, sentence completion and personality inventories, as applied to English adolescents in a secondary modern school. This should be of considerable value in psychological counselling.

On the basis of the Self-Concept Index, the two extreme groups (PSC and NSC) proved to be significantly different from each other on a number of variables, concerned with mental health and good adjustment *vs.* anxiety. The results support Rogers's contention that positive and accepting Self-Concepts are associated with healthy psychological adjustment.

Subjects holding Positive Self-Concepts tend to score higher than subjects holding Negative Self-Concepts and in comparison with Normal Controls on measures of: Psychological adjustment, self esteem, ego and super-ego strength, emotional stability, security, mental health, humanitarianism, non-prejudiced non-stereotyped thinking, perception of parents, boys and girls as valuable, active and potent; social perception status, cognitive clarity, acceptance of Self, perception of school, England, and the World as valuable, literature interest, and academic achievement in Science. Only scores on Authoritarianism pointed to the opposite of the predicted direction.

Care is needed in dealing with sex-associated variables, since sex differences are sometimes so large that one cannot defend mixing the sexes in Self-Concept studies. Girls score higher than boys on some intellectual variables, school evaulation, social adaptation, sociability, premsia, dependence, and femininity, social display interest, liking group action, humanitarianisn and humanitarian interest, variety of interests, sensitivity to concerns of others, non-prejudiced non-stereotyped thinking, tolerant opinion, and evaluation of girls. Boys on the other hand, score higher than girls on mental health, self-confidence, self-sufficiency, self-security, ego strength, emotional stability, dominance, hardness, realism, spatial ability, field independence, rural practical, entertainment, physical, and scientific interests, academic achievement in Science, and boys activity, and future job, potency.

The results substantiate the hypothesis that the Self-Concept is positively and significantly correlated with psychological adjustment, mental health, self-esteem, self-acceptance, acceptance of and by others, ego and super-ego strength, self-sufficiency, academic capability, non-prejudiced non-stereotyped thinking, cognitive clarity, parents and family life evaluation and evaluation of environmental concepts. More specifically, the commonly accepted contention that the individual's Self-Concept is of great inportance for an understanding of the dymanics of personality and adjustment is upheld.

Moreover, it was found that the Self-Concept Index is correlated with lack of scientific interest, sensitivity to concerns of other, evaluation of parents and family life, and liking group action among girls; and is correlated with self sufficiency, ego strength, emotional stability, cognitive clarity, and evaluation of home country among boys.

Assessment of conscious and unconscious aspects of the Self seems necessary for a complete understanding of the Self-Concept. Projective techniques may be of value in eliciting certain aspects of the Self-Concept, which are unobtainable by other techniques, particularly in dealing with maladjusted individuals. But, while it is possible, and advisable to depend on phenomenal approaches, it is risky and insufficient to depend only on non-pheonmenal approaches.

Methodologically, the present study supports the usefulness of the self-rating technique, but recommends the use of some other techniques along with it. The findings, to a large extent, uphold Rogers's belief that the best vantage for understanding behaviour is from the internal frame of reference of the behaver himself.

The results support the hypothesis that in describing the Self-Concept, one cannot simply reverse the characteristics of the Positive Self-Concept and take it as applying to the Negative Self-Concept. Certain variables were found to be particularly characteristic of the Negative Self-Concept, such as Self-devaluation, devaluation of environmental concepts, primitiveness of body concept, sociometric rejection, and lack of interests.

At the same time, a number of variables were found to be unrelated to the Self-Concept. These included intelligence, extraversion, specific interests, masculinity/femininity, and most concepts on the excitability factor in the Semantic Differential.

The use of a Self-Concept Index as a basis for breaking samples down into extreme groups of positive and negative holders is recommended as more adequate than just depending on one single test.

A single positive-negative dimension, if based on a combination of several instruments, represents the major component of the Self-Concept which is relevant to personal adjustment. However, the factor analysis of the battery of tests used showed that the Self-Concept should be regarded as multi-dimensional, covering a number of other social, personality and ability areas.

The Social Desirability factor should be allowed for in self-rating techniques and in clinical situations where clients may be expected, at least at the early sessions, to report on themselves in socially desirable and acceptable directions.

It did not appear to play a large part in the responses of our PSC and NC subjects, but it may produce more distortion among maladjusted subjects.

The results point not only to the importance of Self, but also to the extreme importance of the home environment and the social environment. Environmental treatment should go hand-in-hand with counselling starting from the nearest environment, home, to the wider society, etc. It was hypothesised, and confirmed, that different concepts in the Semantic Differential would range in terms of evaluation from top to bottom as follows: Family concepts, Self concepts, social concepts, environmental concepts, and school concepts. The findings also show that acceptance of Self is positively and significantly correlated with acceptance of, and by, others. They also confirm the belief that the formulation of the Self-Concept is influenced by social factors such as the attitudes of others towards the individual which are usually reflected in his attitudes towards others.

The findings showed that positive Self-Concept and adjustment are significantly related to each other. So are intellectual and perceptual variables. Social and environmental variables are also linked. Interests, which are very important in vocational guidance, failed to gain a comparable status in psychological guidance.

ACKNOWLEDGEMENTS.—Acknowledgements are due to Professor Philip E. Vernon, Professor B. M. Foss, Dr. I. Macfarlane Smith, Dr. P. H. Sandall, Dr. R. Pitts, Dr. P. Slater, Mr. J. E. Kane, Mr. J. Binner, Mr. A. B. H. Brigden, and Mr. T. R Holland.

I am grateful for the help I have received from the staff members of the National Foundation for Educational Research in England and Wales, the University of London and the Institute of Education, and the University of London Institute of Computer Science.

I am indebted to my wife for her regular encouragement and great help.

I should like to express my sincere appreciation to the Government of the U.A.R. for the grant which has made this research possible.

VI.—REFERENCES.

Adorno, T. W., Frenkel-Brunswick ,E., Levinson, D , and Sanford, R. N. (1950). *The Authoritarian Personality.* New York : Harper.

Berdie, R. F., and Layton, W. L. (1957). *Minnesota Counseling Inventory : Manual.* New York : Psychological Corporation.

Burt, C. L. (1937). Correlations between persons. *Brit. J. Psychol.*, 28, 59-96.

Cattell, R. B , and Beloff, H. (1962). *Handbook for the Jr.-Sr. High School Personality Questionnaire.* Champaign, Ill. : Institute for Personality and Ability Testing.

El-Koussy, A. A. H. (1935). An investigation into the factors in tests involving the visual perception of space. *Brit. J. Psychol. Monogr* , Suppl. No. 20, pp. 89.

Gottschaldt, K. (1926). Über den Einfluss der Erfahrung auf die Wahrnehmung von Figuren, I., Über den Einfluss Gehaufter Einpragung von Figuren auf ihre Sichtbarkeit in Umfassenden Konfigurationen. *Psychol. Forsch.*, 8, 261-317.

Naar, R. (1966). Client-centered therapy : a theory of interpersonal relationships. *Mental Hyg.*, 50, 146-149.

Osgood, C. E., Suci, G. J., and Tannenbaum, P H. (1957). *The Measurement of Meaning.* Urbana, Ill. . Univ. of Illinois Press.

Parnell, R. W. (1958) *Behaviour and Physique : An Introduction to Practical and Applied Somatometry.* London Arnold.

Patterson, C. H. (1961). The self in recent Rogerian theory. *Psychologia*, 4, 156-162.

Rogers, C. R. (1942). *Counseling and Psychotherapy, New Concepts in Practice.* Boston Houghton.

Rogers, C. R. (1951). *Client-Centered Therapy : Its Current Practice, Implications and Theory.* Boston, Mass. : Houghton Mifflin.

ROGERS, C. R. (1959). A theory of therapy, personality and interpersonal relationships, as developed in the client-centered framework In KOCH, S. (Ed.), *Psychology: A Study of Science, Vol. III—Formulations of the Person and the Social Context*, pp. 184-256. New York: McGraw-Hill.

ROGERS, C. R. (1961). *On Becoming a Person: A Therapist's View of Psychotherapy.* Boston, Mass.: Houghton Mifflin.

SANDALL, P. H. (1960). *An Analysis of the Interests of Secondary School Pupils: A Factorial Study Leading to the Construction of a Diagnostic Test* Ph.D. Thesis, University of London.

SLATER, P. (1944). *Selective Vocabulary Test.* London: Harrap.

SMITH, I. MACFARLANE (1959). *Manual of Instructions for Spatial Test* 1 (Revised Ed.). London: N.F.E.R.

SMITH, I. MACFARLANE (1964). *Spatial Ability: Its Educational and Social Significance.* London: Univ. of London Press.

STEPHENSON, W. (1953). *The Study of Behavior: Q-Technique and its Metodology.* Chicago: Univ. of Chicago Press.

TANNER, J. M. (1964). *The Pyskique of the Olympic Athlete.* (A study of 137 track and field athletes at the XVIIth Olympic Games Rome, 1960, and a comparison with weight-lifters and wrestlers). London: George Allen and Unwin.

VERNON, P. E. (1963). *Personality Assessment: A Critical Survey.* London: Methuen.

VERNON, P. E. (1965). Envirionmental handicaps and intellectual development, Part II. *Brit. J. Educ. Psychol.*, 35, 117-126.

VERNON, P. E., and PARRY, J. B. (1949). *Personnel Selection in the British Forces.* London: Univ. of London Press.

WITKIN, H. A. (1950). Individual differences in the ease of perception of embedded figures. *J. Person.*, 19, 1-15.

WITKIN, H. A., DYK, R. B., FATERSON, H. F., GOODENOUGH, D. R., and KARP, S. A. (1962). *Psychological Differentiation: Studies of Development.* New York: Wiley.

WRENN, C. G. (1958). The self-concept in counseling. *J. Counsel Psychol.*, 5, 104-109.

WYLIE, R. C. (1961). *The Self-Concept: A Review of the Literature.* Lincoln: Univ. of Nebraska Press.

ZAHRAN, H. A. S. (1964). *An Investigation of Some Aspects of Personality with Reference to the Psychological Guidance of Blind Children.* M.A. Thesis. University of London.

ZAHRAN, H. A. S. (1965). A study of personality differences between blind and sighted children. *Brit. J. Educ. Psychol.*, 35, 329-338.

ZAHRAN, H. A. S. (1966). *The Self-Concept in Relation to the Psychological Guidance of Adolescents: An Experimental Study.* Ph.D. Thesis. University of London.

EVALUATIONS OF THREE FACTORS THAT MIGHT INFLUENCE RESPONSES TO TESTS OF ACADEMIC ACHIEVEMENT MOTIVATION

By CLAUDE E. BUXTON

(*Yale University*)

SUMMARY. The effects of three manipulations in test conditions were compared for a forced-choice and a Likert-type test of academic achievement motivation. (1) Answering as one's friends would, rather than for one's self, slightly decreases mean score, depending upon the school rather than the type of test, and significantly increases variability. (2) Administration of tests by an authority figure (staff member) rather than an outsider produces no significant change in mean score for either type of test when the written instructions indicate confidentiality of responses (3) Instructions to ' fake ' so as to please a teacher produce large and significant changes in mean score in both kinds of test.

Satisfactory reliabilities of the tests and mostly significant relations to criteria of academic motivation are reported. It was concluded that under the conditions typically used in these studies, the Likert test shows somewhat greater effects of social desirability than the forced-choice test, but that more important than test format were the instructions to answer truthfully, and the promise of confidentiality of responses.

I.—BACKGROUND OF THE PROBLEM.

Previous research. I have previously reported (Buxton, 1966) efforts to develop, and compare the validities of, a forced-choice and a Likert-type measure of strength of academic achievement motivation (attitude to attainment). When results on the tests were related to teacher ratings of the strength of academic motivation of secondary school pupils, and to the stream-level of the pupils, the two types of test seemed, somewhat surprisingly, about equally valid. These ways of constructing a test had been compared because it had been assumed that the social desirability of any answer (or the presumed moral evaluation of it by the pupil) would be an important determiner of whether it would be chosen, and the forced-choice test attempts to control for this factor while the Likert test does not. The earlier paper concluded that further examination of the effects of social desirability upon the two types of test was necessary.

It would be difficult to list all the possible social influences upon a pupil's answers about his personal attitudes to any topic, such as school attainment. (For a review of part of the problem, in a discussion of the validity of the forced-choice type of test, see Zavala, 1965). Three of the most immediate and possibly important influences were the object of investigation here : friends' opinions ; presence of an authority figure ; and desire to please the teacher.

II.—ATTITUDES OF FRIENDS.

Procedure. The tests were given in two boys' grammar schools, here called Schools F and G, to totals of 415 and 457 boys, respectively, and were those published previously (Buxton, 1966). The forced-choice test consisted of nineteen paired-sentence choices and the Likert test consisted of the sentence from each of those pairs that had the greater absolute strength-of-motivation rating, disregarding algebraic sign.

The main manipulation in the first study consisted of administering one form of test twice to approximately half the pupils in each school, with differing instructions in the successive administrations. As a boy answered the questions the first time, he was encouraged to give truthful answers pertaining only to himself (here called ' own ' answers) ; as he responded to the test the second time he was asked to give the answer typical of his closest friends. Pupils were promised the results would be confidential to the investigator alone.

Results. (i) General information about the tests. Despite the changes in test scores which are to be described, the successive test administrations showed some similarity in results. The correlations for School F were : for the forced-choice test, ·45, for the Likert test, ·52. For School G the comparable correlations were ·36 and ·39. Reliability coefficients have not previously been reported for the two types of test. Here, when pupils answered according to their own attitudes, the odd-even reliability coefficients, adjusted by the Spearman-Brown prophecy formula, were ·79 in School F for the forced-choice test and ·83 for the Likert test, while in School G the comparable coefficients were ·80 and ·82. For the ' friends ' answers, in School F the reliability coefficients were ·85 for the forced-choice test and ·75 for the Likert test, whereas in School G, the comparable figures were ·88 and ·85. The test responses thus seem stable enough for purposes of group comparisons.

(ii) Changes in test score distributions. Shown in the upper part of Table 1 are the data of main interest. The consistent observation is that the mean score for academic achievement motivation decreases when pupils are asked to answer as their friends would. However, the decrease is statistically significant for both types of test in School F, and for neither type of test in School G. Such a finding underlines my having said in the earlier paper that differences between school populations may sometimes out-weigh other causes of differences in test results. It is to be noted that changes in mean score are small, i.e., pupils do, in fact, have attitudes rather like those of their friends, yet the social desirability of answers, despite this, may have caused the ' own ' scores to be elevated slightly.

TABLE 1

COMPARISONS OF MEANS, VARIANCES AND TEST TYPES FOR ' OWN ' and ' FRIENDS ' CONDITIONS

Tests of Mean Differences							
School, Test	M_{own}	S.D.$_{own}$	$M_{fr.}$	S.D.$_{fr.}$	t	d.f.	P
F, forced choice	11·4	4·0	9·0	4·6	7·32	205	< ·01
G, forced choice	10·9	4·1	10·4	5·2	< 1·00	227	—
F, Likert	71·9	10·7	63·7	15·1	8·00	208	< ·01
G, Likert	70·4	9·9	67·8	18·2	1·06	228	—
	Tests of Differences in Variance*						
F, forced choice					2·07	204	< ·05
G, forced choice					3·95	226	< ·01
F, Likert					6·40	207	< ·01
G, Likert					10·39	227	< ·01

* The test used here, for correlated data, is provided by formula 8.7 of Walker and Lev (1953).

The lower part of Table 1 shows an unexpected finding which can be explained only in part: the variability of scores on each type of test increased significantly in all four comparisons. To analyze this finding futher, the eight frequency distributions in the four comparisons were plotted, and it is necessary to comment on each comparison separately. In School F, where mean forced-choice score decreased significantly, there was some skewness in the 'own' condition toward making one's self 'look better' than one's friends. The skewness was slightly reversed in the 'friends' condition and the distribution became flatter, suggesting the possibility of two factors at work: social desirability, which decreased in the 'friends' condition, and also a stereotyped view about what one's friends' answers might be, so that in this condition there was a rather consistent tendency for some pupils to go to one extreme or the other in choosing answers. How important either factor is cannot be estimated. In School G, where the forced-choice mean did not change significantly from 'own' to 'friends,' there was a rather conspicuous flattening of the distribution in the second condition, so that again and more obviously, extreme scores occurred and the idea of scores moving to the extremes of the scale because of a stereotyped idea of 'friends' attitudes is again plausible.

For the Likert test the results appear to have a more obvious explanation. In both School F, where the mean score decreased significantly in the 'friends' condition, and in School G where the decrease was insignificant, the frequency distributions show the same result. In both schools, the 'own' distributions had been quite skewed in favor of making one's self 'look good'; in the 'friends' condition the skewing was much less marked and there was room for greater spread of scores at the upper end of the scale.

In view of the fact that the amount of skew in the forced-choice distributions is not great, whereas it is somewhat greater in the Likert distributions, we may tentatively conclude that the expected influence of social desirability is showing itself here. The increased variabilities may then in a sense be artifacts of this. The decrease in the mean scores for one school in the 'friends' condition, and the insignificance of the same change in the other school, suggest that the extent to which pupils wish to make themselves 'look better' than their friends is a function of variables not identified here.

(iii) Relations of test scores to criteria of academic motivation. In my previous paper (Buxton, 1966), the main concern was for relative validities of the two types of test here employed, and so rather extensive tabular evidence was provided in making comparisons. All the same evidence is available here, but it will be reported only briefly.

In School F analyses of variance showed that both the forced-choice test and the Likert test were significantly related ($P < \cdot 01$) to teacher ratings of pupils when 'own' answers were given, thus extending the cross-validations previously reported. When 'friends' answers were given, the level of significance dropped to the ·05 level for the forced-choice test and just missed this level for the Likert test. Nevertheless, the relationships remained clear.

In School G the 'own' scores related significantly ($P < \cdot 01$) to teacher ratings for the forced-choice test and barely missed the ·05 level for the Likert test. Under the 'friends' instruction both relationships became insignificant. Again, the 'own' results tend to agree with earlier cross-validations, but in this school the 'friends' results imply that a good many pupils were probably reacting as though their friends were consistently different from themselves.

In School F neither the forced-choice nor the Likert tests related significantly to the stream level of the pupil under either instruction condition. In School G, however, both types of test were significantly related ($P < \cdot 01$) to stream level under the 'own' instruction whereas under the 'friends' instruction the forced-choice relation became insignificant and the Likert-test relation weakened somewhat but was still significant ($P < \cdot 05$). The cross-validation under 'own' instructions only partially supports previous results, and again we can only point to unspecified differences between schools.

Conclusions. These data show only a few clear consequences of instructing that 'friends' answers be given rather than 'own'. (*a*) Mean test scores may decrease significantly, or not, depending on the school. (*b*) Variability of response increases conspicuously when pupils answer as they think their friends might. I tend toward somewhat different emphases in interpreting results for the two kinds of test: on the Likert test, it may be fairly important that the pupils tended to make themselves 'look good' when giving 'own' answers, thus producing a distribution which was sufficiently skewed that when they answered as 'friends' the distribution moved away from the upper end of the scale so that more variability was possible; on the forced-choice test, in the 'friends' condition it may be important that the pupils tended to answer according to a stereotype of what their friends would say, thus producing more extreme scores and a flattened distribution. (This is not to argue that there was no stereotype effect confounded in the results for the Likert test.) (*c*) Both types of test fairly successfully predict teacher ratings under 'own' instructions. They less consistently predict stream level.

Thus far, the implication that social desirability or moral evaluation of responses raises scores more in the Likert 'own' test, is about the only suggestive basis for electing to use the forced-choice format instead.

III.—PRESENCE OF AN AUTHORITY FIGURE.

Procedure. The second study centered on the question whether the identity of the person giving the tests influences answers, and if so, whether there is a difference in the amount of effect for forced-choice and Likert tests. The manipulation was a very simple one; half the classes in a school were tested by an outsider (the same person who collected all the data in the other studies reported here). The other classes, equated with the first set as carefully as possible, were given the test by a member of the school staff. This teacher explained his role to the pupils by an off-hand remark indicating that the tests were to be given throughout the school, and since several of the staff were interested in how the results might turn out, he was merely helping out. He took pains to permit the implication that staff members (authority figures) *might* see individual pupil responses to the questionnaire (although, in point of fact, they did not). Half the pupils in each class received a forced-choice test and half a Likert test. In School H, the number of pupils tested was 432, and in School J, 402. Both were boys' grammar schools.

Results. (i) Effects of presence of authority figure. As shown in Table 2, in neither school did the variation in the tester make for a significant change in mean score on either test, although three of the four means increased slightly with staff administration of the tests. This general finding may be the result of the way in which we had resolved a procedural dilemma: the usual written instructions given the pupil promised confidentiality of results; should we, or should we not, remove this promise from the instructions for tests used by either kind of tester? Our resolution was to keep the promise of confidentiality as normally printed, for both testers, so that the off-hand remark by the teacher, and his presence, were the only efforts at manipulation of the authority variable In effect, the pupil would have had to disbelieve the printed instructions to be influenced to respond by making himself 'look good' to an authority figure.

TABLE 2

COMPARISONS OF MEANS AND TEST TYPES FOR ' OUTSIDER ' AND ' AUTHORITY-FIGURE ' CONDITIONS

School, Test	M_{out}	$S.D._{out}$	M_{auth}	$S.D._{auth}$	t	d.f.*	P
H, forced choice	10·1	1·3	10·9	1·2	1·38	118	—
J, forced choice	10·0	1·4	10·6	1·2	1·09	99	—
H, Likert	66·8	11·1	67·4	10·0	<1·00	107	—
J, Likert	69·9	11·6	69·5	8·9	<1·00	102	—

* Mid-number of d.f. for groups of slightly different size in three of four comparisons; groups identical in size in School J, Likert test.

Therefore, the results obtained here may simply be due to a weak experimental manipulation, i.e., the pupil read and trusted what was set before him in the instructions, hence pupils to whom a teacher administered the test behaved no differently than those having the outsider.

(ii) Relations of test scores to criteria of academic motivation. These data can be reported even more briefly than for the preceding study. For the first time, in School H, neither type of test related to either teacher ratings or stream-classification. In School J, however, both the forced-choice and the Likert tests related significantly ($P < \cdot 01$) to both criteria when the outsider gave the tests, but insignificantly in three of four evaluations when the teacher gave the test (the fourth was a significant relation of Likert scores, $P < \cdot 01$, to teacher ratings). These findings suggest that although mean scores did not vary significantly from one kind of tester to another other influences were at work when the teacher gave the test in School J, so that the pupils received scores not relating to the criteria in the usual way. At the risk of redundancy, it must again be said that in School H this implication may not be drawn because of unknown variations pertaining, presumably, to the school as a whole.

IV.—DESIRE TO PLEASE THE TEACHER.

Procedure. The final study of this series was concerned with whether what the teacher expects, as a representative of adult standards for attitudes to school attainment, affects the way pupils respond. The two types of test were, therefore, administered twice to each pupil taking either the forced-choice or the Likert-type test, with the instruction first to answer truthfully (here called ' own ' responses, again), and second, to answer in the way that would please a teacher. A significant change in score, on either type of test, indicates both that it is *possible* for the pupil to ' fake ' to please a teacher and, by implication that, under normal instructions, including the promise of confidentiality of answers, this factor is not very important. A total of 424 boys, half having each type of test, were tested in School K, and 590 boys divided the same way were tested in School L. Both were grammar schools.

Results. (i) General. Odd-even reliabilities, adjusted by the Spearman-Brown prophecy formula, were again computed; they were of the same general order of magnitude for the ' own ' condition as the coefficients reported in Section II of this paper. Under the ' teacher ' condition the coefficients decreased somewhat, probably because of a restriction of range as will be recorded but not discussed later. The correlations between ' own ' scores and ' teacher ' scores approximated zero for both types of test in both schools.

(ii) Changes in test score distributions. Shown in Table 3 are the data of main interest. It is very clear that the manipulation performed in this study was effective : the pupils, regardless of type of test or school, could give answers that would be called correct or desirable by a teacher, and each of the four comparisons demonstrates this at a high level of significance. The frequency distributions are not shown here, but one can readily infer from the relations between means and possible ranges of scores what happened with respect to the shapes of the distributions. The forced-choice test scores, as typical in all these studies where the pupil was asked to give ' own ' answers which would be kept confidential, are only slightly skewed to as to make the pupils ' look better ' than they would if the distributions were symmetrical. But when the pupils are asked to give answers that would please a teacher, there is an extreme shift in the distribution so that most scores are near the highest possible score, 19. The effect is also present in the Likert test. On this test, there is typically somewhat more skew than on the forced-choice test in the ' own ' confidential mode of response (often the means are near 70 within a range of 5–95) ; consequently, although the average score under instruction to please a teacher is fairly close to the maximum, the relative shift can not be so great in the Likert scores as in the forced-choice scores.

TABLE 3

COMPARIONS OF MEAN SCORES FOR ' OWN ' AND ' TEACHER ' CONDITIONS.

School, Test	M_{own}	$S.D._{own}$	M_{tea}	$S.D._{tea}$	t	d.f.	P
K, forced choice	11·8	3·8	17·1	1·8	18·28	211	$<\cdot01$
L, forced choice	10·7	4·1	17·2	1·6	26·15	295	$<\cdot01$
K, Likert . . .	68·8	9·6	84·6	6·2	21·2	211	$<\cdot01$
L, Likert	68·1	9·4	85·5	5·7	27·9	295	$<\cdot01$

(iii) Relations of test scores to criteria of academic motivation. Analyses of variance relating teacher rating to test scores were again computed for both types of test and both schools. For the ' own ' condition three F values were highly significant ($P < \cdot01$) and one, for the forced-choice test in School K, was well beyond the ·05 level. F values computed for the ' teacher ' condition were all small and insignificant. In comparable analyses of variance relating stream-level of the pupil to test score, only one F (School K, ' own ' condition, Likert test) was significant ($P < \cdot01$).

Conclusion. When asked to do so, pupils clearly can give the answers, on either type of test, that would please a teacher, i.e., show high motivation. Conversely the results suggest that, under the usual conditions of administration of the two types of test, we have been investigating in these studies (truthfulness requested, confidentiality promised), desire to please the teacher, or perhaps desire to make one's self ' look good ' by adult standards, is not a very important influence. This is, perhaps, less true of the Likert-type test, as shown by the greater degree of skewness in most of the studies.

The Likert test has shown somewhat greater variability than the forced-choice test in its relations with criteria of strength of academic motivation, especially with the criterion of stream-level (where it more than once surpassed the forced-choice test). This may be a function of the kind of test it is, but one begins to wonder also whether stream-level, like marks earned (see Buxton, 1966), is defined in such different ways from school to school that one should doubt whether it is a very useful criterion.

V.—DISCUSSION AND CONCLUSIONS.

Taken together, these three studies do not differentiate sharply between the virtues of forced-choice and Likert tests for the purpose to which they have been put here, namely, measuring strength of academic motivation in secondary school boys asked to tell the truth under promise of confidentiality. The indications are quite clear that the forced-choice test does not effectively conceal which is the socially desirable answer, any more than does the Likert test. It is also fairly clear that there is an upward trend in Likert scores under ' own ' conditions, as though the format of the test encourages a ' look good ' response whether the pupil intends it or not. Instructions to answer as one's friends might produce significant but small decreases in score, and administration of the test by a person in authority over the pupil who has, nevertheless, been promised confidentiality of results produces no significant score changes. But instructions to give a ' good ' answer are very effective.

From all this we can probably best conclude what type of test is of some importance under the conditions employed here (the forced-choice type being slightly favored), and might be of greater importance under other conditions such as individual diagnostic measurement. But, in these studies, the intention or set of the pupil to give truthful answers about himself in confidence seems to be more important. (The conclusion that social desirability response sets are relatively unimportant for their junior-level personality test has also been reached with a comparable procedure and similar reasoning by Eysenck, *et al*, 1966.

Two additional comments seem required. The first is that the social desirability ratings of the sentences used in the tests reported here are now several years old ; presumably social desirability is a changing factor among school pupils and new ratings or new statements are needed for future work. The second remark is that social desirability has been shown by Wiggins (1966) to be complex or multi-dimensional (and some of my own recent unpublished work shows the same result.). In the future, the construction of a forced-choice test utilizing ratings of social desirability is, therefore, going to be even more complex than in the past in any domain of personality or attitude.

ACKNOWLEDGEMENTS.—Mr. Frederick G. S. Turner was most helpful in gathering the data, and Mr. Charles M. Super in analyzing them on the electronic computers at Yale. The Kettering Foundation provided financial support, for which I am most grateful. Mr. Turner and I am much in debt to the Headmasters and staffs of the following schools, all in the greater London area : Beal Grammar School ; Cooper's Company Grammar School ; Ealing Grammar School ; Wallington County Boys' Grammar School ; St. Nicholas' Grammar School ; Christ's College.

VI.—REFERENCES.

BUXTON, C. E (1966). Evaluations of forced-choice and Likert-type tests of motivation to academic achievement. *Brit. J. Educ. Psychol.*, 36, 192-201.

EYSENCK, S. B. G., SYED, I. A., and EYSENCK, H. J. (1966). Desirability response set in children. *Brit J. Educ. Psychol.*, 36, 87-94.

WALKER, H. M., and LEV, J. (1953). *Statistical Inference.* New York : Holt, Rinehart and Winston.

WIGGINS, N. (1966). Individual viewpoints of social desirability. *Psychol. Bull.*, 66 68-77.

ZAVALA, A. (1965). Development of the forced-choice rating scale technique. *Psychol. Bull*, 63, 117-124.

(*Manuscript received 18th November*, 1966)

RESEARCH NOTES

THE EFFECT OF RESTRICTION OF RANGE UPON CORRELATION COEFFICIENTS

By STEPHEN WISEMAN

(*University of Manchester*: *School of Education*)

SUMMARY. A recent paper by Swift reports correlations between social class and various educational measures, and shows that restricting the range of the social class measure may increase the correlation above that for the whole range. This is contrary to the general expectation of shrinkage described in the statistics text books. This paper shows that this is far from an isolated case, and lists some other examples. The relevance of this phenomenon to experimental design in educational psychology is discussed.

One of the 'facts of life' which all students of elementary statistics have drilled into them is that restricting the range of ability in a sample produces a shrinkage of correlations. As Peters and Van Voorhis (1940) say, "It is well known that, other things being equal, the size of a coffiecient of correlation is very much affected by the heterogeneity of the population on which it is computed" (p. 208) According to Guilford (1942), "The size of *r* is very much dependent upon the range of ability or, in more general terms, the variability of measured values, in the correlated sample. The greater the variability, the higher will be the correlation, everything else being equal" (pp. 248-9). We are indebted to Karl Pearson (1903) for formulae which enable us to compute what the correlation would be in the unrestricted group, provided that we know the standard deviation of at least one of the variables in both the restricted and the unrestricted group:

$$R_{12} = \sqrt{1-\frac{s_2^2}{S_2^2}\left(1-r_{12}^2\right)}$$ *

A paper by Swift (1967) in the last issue of this *Journal* reports an analysis of of the relationship between social class and 11+ success, in which it is shown that restricting the range of social class may result in *increasing* the correlation. He concludes that "it is better to treat social classes as social contexts within which particular factors may have special effects upon cognition and motivation. Analysis of sample data suggests that the effects of certain environmental variables are different in the different social classes." It is the purpose of this paper, first, to suggest, not that Swift's conclusion is erroneous, but that it is not the only interpretation possible; second, to show that the statistical phenomenon reported by Swift is far from being a special case, and that many other examples exist.

EXAMPLES FROM MANCHESTER SURVEYS.

The Manchester researches on environment and education reported in Wiseman (1964) showed significant differences in inter-school analyses for bright children and for backward children. We have recently engaged on a more detailed survey of primary school children (Wiseman 1967) in which similar differences occurred. Since the writing of that report, further analysis of the data has been carried out, and this has revealed several striking examples of the phenomenon noted above. A representative sub-sample of 186 children of 10+ from twenty-two primary schools formed the basis of this further analysis. For these children we had test scores for verbal intelligence, English and arithmetic, together with very many measures of environmental variables covering home, school and neighbourhood.

---

* Thorndike (1949), p. 173. He also gives formulae for other cases: e g, where curtailment has been on the basis of a third variable.

Many of the variables derived from interviews with the parents, carried out by the Social Survey of the C.O.I.

In addition to the normal correlation coefficients computed between the various tests and the environmental or social variables, we obtained two other sets of correlations. The sample of 186 children was split at the median verbal I.Q. score, and correlations calculated for the ' high ' group and the ' low ' group separately. One would expect—the following the statistics text books—that the correlations for these two drastically curtailed samples would be significantly smaller than those from the full range of ability. With many of the social variables, however, this was far from the case, as may be seen from Table 1.

TABLE 1

CORRELATIONS OF TEST-SCORE WITH SOCIAL VARIABLES : CURTAILMENT ON I.Q.

Variables correlated	Full Range N=186	High Group N=93	Low Group N=93
Father's occupation .			
Arithmetic	·000	—·274	—·004
English	—·121	—·220	+·023
Verbal Intelligence	—·088	—·119	—·007
Parental talks with teacher :			
Arithmetic	+·148	+·095	—·206
English	+·155	+·274	—·125
Verbal Intelligence	+·189	+·162	—·114
Father's interest in school work :			
Arithmetic	+·125	+·205	—·134
English	+·127	+·229	—·179
Verbal Intelligence	+·176	+·260	—·050

These results are even more dramatic than those of Swift. Not only do we have an *increase* of correlation when we consider only one-half of the intelligence range, but in many cases we have a reversal in sign. And the Table is merely a selection from a much larger number of variables, many of which display the same kind of differences.

The first three lines of Table 1 provide very similar data to those of Swift—except that in our case we have restricted the range on intellectual ability instead of on social class. We could, perhaps, paraphrase Swift's conclusion and say " it is better to treat intellectual levels as intellectual contexts within which particular factors may have special effects upon social behaviour and attitude."

EXAMPLES FROM OTHER FIELDS.

So far, our examples have come from the field of educational /social interaction. But investigation shows that this phenomenon is far from restricted to this sphere of relationships. A follow-up of college of education students, reported in Wiseman and Start (1965) has been carried further by Start (1966) with a more intensive analysis. One of the criterion measures of teaching success in the school five years after training was a rating by the head teacher. In a sample of 450 teachers, a split was made at the median of the head's rating, producing two groups of 225, one ' high ' and one ' low ' Correlations within each of these groups showed some interesting differences, as may be seen from Table 2. The first line of the table shows the correlation between the head's rating and the measure of success computed from the aggregate of four judges assessing the head teacher's confidential reference. This shows the ' expected ' shrinkage from the total group to the sub groups as postulated by the textbooks (although even here the difference between ' high ' and ' low ' is far from negligible). The following lines of the Table, however, are in sharp contrast to this.

TABLE 2

CORRELATIONS FROM TEACHER FOLLOW-UP : CURTAILMENT ON HEAD'S RATING OF QUALITY OF TEACHER

Variables Correlated		Full Range N=450	High Group N=225	Low Group N=225
Head's Rating	Judge's Rating	·781	·336	·570
Special Responsibility allowance	Dissatisfaction with teaching	·032	·188	—·065
No. of courses attended	Judges' rating	·113	·135	—·101
College teaching mark	Head's rating	·250	·050	·230
	Maths. content of main subject	·031	·190	—·138
College total academic score	Maths. content of main subject	·228	·344	·095
Principles of education mark	Maths. content of main subject	—·086	·018	—·194
College total academic score	Is teaching subject trained for	·149	·198	—·087

Yet another example is given in Table 3. This is taken from White (1965), who investigated the role of interest in the selection of craft apprentices.

TABLE 3

CORRELATIONS WITH PRACTICAL ASSESSMENT : CURTAILMENT ON MOTIVATION TEST SCORE (INTEREST)

Variable	Total Group	Interest grade from motivation test				
		A	B	C	D	E
Descriptive practical interest	·300	·395	·462	·449	—·054	·191
Practical Knowledge	·558	·700	·593	·541	·444	·523
General Ability	·340	·293	·266	·258	·402	·429
Spatial	·406	·216	·452	·410	·353	·553
Mechanical Comp.	·410	·348	·483	·472	·271	·433
Arithmetic	·179	·307	·436	·093	·165	·300
English	·371	·262	·273	·400	·375	·590
Clerical Aptitude	·368	·540	·299	·186	·407	·494
N	234	45	55	50	54	30

Here we have a more detailed analysis, since White divides his population into five sub samples on the basis of his test of motivation, sub group A representing those subjects whose occupational motivation is predominantly that of interest in the work, as against factors of working conditions, wage-level, career-prospects, etc. Succeeding groups represent lower and lower levels of weight given to interest as a motivator. The traditional shrinkage of correlations caused by such a radical reduction in heterogeneity is almost completely absent. White's main thesis—that an interest test is a useful predictor of technical competence *if* one restricts its use to those subjects who rate interest high as a motivator—was clearly proved, for the correlation of his interest tests with G.C.E. marks in woodwork and metalwork was ·300 in the total sample, but ·428 when only groups A, B and C were included.

The list of examples is by no means exhausted. In a comprehensive article, Eysenck (1966) cites a large number of investigations in the field of personality in which sub-samples of the population studied show either a higher level of association in particular variables than the parent population, or display associations in opposite directions. He suggests " that correlational psychology, just like experimental psychology, is subject to typological troubles, and erroneously makes the same assumptions of lack of individual differences—even when investigating individual differences ! . . . in running correlations over random samples of the population, the (implicit) assumption is made (in factor analysis) that the same system of relations obtains in all members of this population ; in other words, individual differences in structure of abilities, or temperamental traits, is denied. Yet it can be shown that they exist ; the structure of intellectual abilities is different in stable and in labile children, and the structure of personality traits in bright and in dull students " (p. 23). Many of Eysenck's examples are, as he makes clear, cases of non-linear regression, but it should be noted that it is quite possible—and even common—for cases exhibiting linear regression to produce higher correlations in part than in whole. If the scatter diagram is triangular instead of ellipsoidal, then the correlation at ' the sharp end ' will be higher than over the whole range, even though the regression line is undoubtedly straight.

DISCUSSION.

It will be noted that all the examples cited in this paper involve variables within the field of personality, attitude and motivation (if we assume, as perhaps we may, that the effective elements in, if not the determinants of, ' social class ' are very largely of this nature). It seems less likely that the phenomenon operates in correlations between purely cognitive variables—but we cannot be sure. The lack of obvious examples suggests that, at any rate, the mechanism is much rarer in this restricted field. But this is cold comfort, since nowadays the great majority of worthwhile investigations in education and educational psychology must involve measures of motivation, attitude and personality.

Since the development first of electric desk calculators and second of electronic computers, the danger of drawing false conclusions or reaching only partial or misleading ' insights ' by the lack of recognition of the operation of this phenomenon has risen sharply. When we had to do all our work ' by hand ' we were restricted to relatively small samples and a dozen or so variables. The existence of odd results in even smaller sub-samples, even if recognised, could reasonably be shrugged off as idiosyncrasies produced by the sampling process. With larger and larger samples, and larger and larger matrices, the data go straight on to punched cards or tape, and the experimenter never sees the shape of the correlation scatter—and this is the major (perhaps the only) warning signal.

What conclusions are we to draw ? There are two major ones, perhaps. First, that our statistics text-books for students of educational psychology must include clear warnings about the limitations of their ' restriction of range ' sections, and cease to suggest that correlation shrinkage due to this cause is well-nigh universal. Second, in addition to being alert to the possibility of this effect occurring in *any* research, it ought to become almost routine for experimenters in education to divide their samples into top, middle and bottom thirds on the basis of intellectual ability (or academic ability) and to do separate correlational analyses on each of these sub-samples independently. The present wave of interest in educational sociology, and the many enquiries proceeding in the field of school organisation—streaming and setting, the different forms of ' comprehensive ' structure—may produce some highly misleading results unless the researchers are aware of this particular phenomenon and its implications.

REFERENCES.

EYSENCK, H. J. (1966). Personality and experimental psychology. *Bull. of Brit. Psych. Soc.*, 19, 62, 1-28.

GUILFORD, J. P. (1942). *Fundamental Statistics in Psychology and Education* New York : McGraw Hill.

PEARSON, K (1903). Mathematical contributions to the theory of evolution—XI. On the influence of Natural Selection on the variability and correlation of organs *Trans. Royal Soc* (*London*), Series A, 200, 1-66.

PETERS, C. C., and VAN VOORHIS, W. R. (1940). *Statistical Procedures and their Mathematical Bases*. New York . McGraw Hill

START, K. B. (1966). The relation of teaching ability to measures of personality *Brit. J.* Educ Psychol, 36, 158-165.

SWIFT, D. F. (1967). Family environment and 11+ success. *Brit J. Educ Psychol.*, 37, 10-21.

THORNDIKE, R. L. (1949). *Personnel Selection Test and Measurement Techniques.* New York : John Wiley.

WHITE, J. E. (1965). *The Role of Interest in the Selection of Craft Apprentices.* M.Ed. Thesis Univ. of Manchester Library.

WISEMAN, S. (1964). *Education and Environment* Manchester : Manch. Univ. Press.

WISEMAN, S. (1967). The Manchester Survey. *Report of the Central Advisory Council for Education.* (The Plowden Report). Appendix 9, 347-400.

WISEMAN, S., and START, K. B. (1965). A follow-up of teachers, five years after completing their training. *Brit. J. Educ. Psychol.*, 35, 342-361.

PREDICTING READING AGES ON NEALE'S ANALYSIS OF READING ABILITY

BY WILLIAM YULE

(*University of London Institute of Education*)

I.—INTRODUCTION.

In assessing degree of reading retardation, it is not sufficient to consider merely the discrepancy between the child's chronological age and his reading age on some standard measure. It is desirable and necessary to consider also the child's intelligence. Among the most recent studies attempting to do this are those by Ravenette (1961), and Fransella and Gerver (1965). Both these studies used Schonell's Graded Word Reading Test (1950) as their measure of reading attainment. However, the former study did not take chronological age into account in predicting reading age, and the latter study relied on data gathered in the children's department of a psychiatric hospital. Thus, we still do not know what size of discrepancy to expect in a normal sample of children when reading age is predicted from both intelligence and chronological age.

In any case, as Curr and Gourlay (1960) have pointed out in their study of the effect of repeated exposure to Schonell's G.W.R.T., the usefulness of this test is seriously undermined in clinics which have no control over the child's previous reading-test experience. Since the Schonell G.W.R.T. is widely used in the classroom, it would be advisable not to use it in clinics. An alternative test, which is rapidly becoming more widely used, is Neale's Analysis of Reading Ability (1958). This was designed specifically as a diagnostic test and much useful qualitative information can be gathered at the same time as three separate quantitative measures of reading ability—a measure of the Rate of Reading, one of the Accuracy of reading and a third of the Comprehension of Reading. Since the Schonell test would seem to be primarily one of accuracy of reading, Neale's analysis immediately offers certain advantages which outweigh the disadvantage of the slightly longer time taken to administer it. Furthermore, as Netley, *et al* (1965) have demonstrated, Neale's Test does not suffer from repeated exposure and is, therefore, admirably suitable for retesting to assess progress.

II.—THE PRESENT STUDY.

One-hundred-and-sixty children between 9 and 11 years of age were chosen at random from the total population of some 2,200 children in this age range whose parents resided on the Isle of Wight. As reported in detail elsewhere (Rutter, *et al*, 1967) this group was found to be representative of the general population, even after the thirteen children whose parents refused permission for further testing were omitted from the intensive study.

Included in the large battery of tests administered to this control group, were the Neale's Analysis of Reading Ability described above, and a short form of the WISC based in Maxwell's (1959) factor analysis of Wechsler's original data. The short form consists of four subtests : Similarities, Vocabulary, Block Design and Object Assembly. The sum of the scale scores on these four sub-tests acts as our measure of I.Q. The means and standard deviations of the measures used are shown in Table 1. It is interesting to note in passing that unlike Netley's sample in South London, this sample is not retarded on any of the reading measures. The mean sum of the WISC scale scores of 46 is equivalent to a Full Scale I.Q. of 107, in Wechsler terms, and can be interpreted as a reflection of Jones' (1962) finding that British children tend to score above the American norms on this test.

TABLE 1

MEANS AND STANDARD DEVIATIONS OF RELEVANT VARIABLES (N=147)

	Mean	S.D.
Chronological Age in Months	121·13	6·90
Rate of Reading in Months	121·86	20·87
Accuracy of Reading in Months	121·98	18·00
Comprehension of Reading in Months	125·48	18·35
Sum of WISC Scale Scores	45·99	8·71

III.—RESULTS.

The multiple regression equations are shown in Table 2 The standard errors of estimate are given in months and these compare very favourably with those of previous studies. From the data it is possible to say what size of discrepancy between measured and predicted reading ages occur in various percentages of the standardization sample at each end of the reading continuum. In studying retardation, we are concerned with only one tail of the distribution, and Table 3 shows the size of discrepancy expected at crucial levels. Thus, approximately 10 per cent. of children can be expected to be twenty months retarded from their predicted accuracy of reading age. Only future studies of children screened in this way will show where the line should be drawn in considering which children to select for remedial instruction, but in the meantime, the above procedure allows us to pick out the most severely retarded who, presumably, can best profit from the services available.

TABLE 2

MULTIPLE REGRESSION EQUATIONS FOR PREDICTING THE THREE READING AGES ON NEALE'S ANALYSIS OF READING.

Measure	Multiple Regression Equation	Standard Error of Estimate
Rate of Reading	$28{\cdot}45+(0{\cdot}96\times\text{“ I.Q.”})+(0{\cdot}41\times CA)$	19·11
Accuracy of Reading . .	$-3{\cdot}87+(0{\cdot}93\times\text{“ I.Q.”})+(0{\cdot}68\times CA)$	15·74
Comprehension of Reading	$-23{\cdot}44+(1{\cdot}15\times\text{“ I.Q.”})+(0{\cdot}79\times CA)$	14·95

" I.Q."—Sum of scaled scores on Similarities, Vocabulary, Block Design and Object Assembly on the WISC.

CA—Chronological Age in Months.

TABLE 3

SIZE OF DISCREPANCIES (IN MONTHS) BETWEEN PREDICTED AND OBSERVED READING AGES BY VARIOUS PERCENTAGES OF THE STANDARDIZATION POPULATION (ONE TAIL TEST).

Measure	20%	10%	5%	1%
Rate	16·05	24·46	31·34	42·62
Accuracy of Reading	13·22	20·15	25·81	35·10
Comprehension of Reading	12·56	19·14	24·52	33·34

TABLE 4

INTERCORRELATIONS OF ACCURACY, COMPREHENSION AND RATE SCORES (N=147)

	Accuracy	Comprehension
Rate	0·75	0·53
Accuracy	—	0·77

The intercorrelations among the three subtests are shown in Table 4. It can be seen that the same patten is present as was found in Netley's study. Again, it is hoped that the information contained in this table will prove useful to clinicians in interpreting subtest discrepancies.

IV.—RETEST DATA ON THE MEASURES.

The results presented above form part of a study carried out over a two-year period. Testing of children screened out for various reasons was undertaken in two periods separated by exactly one year. It was discovered that 140 children aged originally 9 to 11 years had been tested on both occasions on the measures described earlier. Naturally, this data must be treated with caution since all children were screened out of the total population because of suspected educational handicap in the first instance and a year later because of suspected psychiatric or physical handicap. In fact, thirty-three of the 140 were found to have a severe reading retardation defined in terms of a discrepancy of at least twenty-eight months on either the Accuracy or Comprehension Scores and those scores predicted on the basis of their age and intelligence; fifty-eight had significant psychiatric disorders, twenty-five had significant physical disorders, six had both psychiatric and physical disorders, so that seventy-seven of the 140 children had at least one handicapping condition on the second occasion of testing. Even so, it is possible to present some badly needed test-retest information on the WISC, in particular, as well as on the reading test. This is presented in Table 5.

TABLE 5

TEST RE-TEST CORRELATIONS ON A SAMPLE OF 140 CHILDREN RETESTED AFTER AN INTERVAL OF ONE YEAR.

Measure	Rtt.
Rate	0·76
Accuracy	0·95
Comprehension	0·93
Short WISC	0·86

All correlations are significant beyond P=0·001.

It can be seen that the Accuracy and Comprehension Scales of the reading test are remarkably stable, but, as reflected in the larger standard error of estimate in Table 2, the Rate of reading would appear to fluctuate too much to be of much use in individual prediction over a one-year period. The test-retest correlation of 0.86 on the sum of the four subscales of the WISC fully justifies its use as a screening instrument as described in the present study.

V.—Discussion.

Neale's Analysis of Reading Ability has shown itself to be an extremely reliable measure over a one-year period, even with a slightly abnormal sample. It is a simple test to administer. Children find it an interesting, challenging test to attempt, and clinicians can gain useful insight into the pattern of errors commonly made in prose reading. (This latter aspect has yet to be explored fully as a tool for differential diagnosis of reading difficulties). As such, then, it is admirably suited to screening out children with significant discrepancies (empirically defined) between expected and actual reading attainment.

The short form of the WISC used as the measure of intelligence has also been shown to be reasonably reliable over a one-year period. Where a full WISC has been given, only the sum of the four scale scores mentioned should be entered into the prediction equation.

It is hoped that the above data, based on a random and representative sample of children aged 9 to 11 years, will prove a useful basis for interpreting the degree of reading retardation exhibited on testing.

Acknowledgements.—This study forms part of a series of epidemiological enquiries concerning educational, physical and psychiatric handicaps among school-age children, who are resident on the Isle of Wight, carried out in collaboration with Prof. J. Tizard of the University of London Institute of Education, Dr. M. Rutter and Dr. P. Graham of the Institute of Psychiatry, and Dr. K. Whitmore of the Department of Education and Science. Throughout these enquiries we have enjoyed the whole-hearted co-operation of the parents, the schools doctors and local authorities on the Island. We are particularly grateful to Dr. R. K. Machell, Principal School Medical Officer, to Mr. A. L. Hutchinson, County Education Officer, and to Mr. L. Rigley, Senior Research Officer, without whose help these studies would not have been possible. The study was financed in part by the Department of Education and Science, the American Association for the Aid of Crippled Children, and the Medical Research Council.

VI.—References.

Curr, W., and Gourlay, N. (1960). The effect of practice on performance in scholastic tests. *Brit. J. Educ. Psychol.*, 30, 156-167.

Fransella, F., and Gerver, D. (1965). Multiple regression equations for predicting reading age from chronological age and WISC Verbal I.Q. *Brit. J. Educ. Psychol.* 35, 86-89.

Jones, S. (1962). The Wechsler Intelligence Scale for Children applied to a sample of London primary school children. *Brit. J. Educ. Psychol.*, 32, 119-132.

Maxwell, A. E. (1959). A factor analysis of the Wechsler Intelligence Scale for Children. *Brit. J. Educ. Psychol.*, 29, 237-241.

Neale, M. D. (1958). *Neale Analysis of Reading Ability Manual.* London: Macmillan

Netley, C., Rachman, S., and Turner, R. K. (1965). The effect of practice on performance in a reading attainment test. *Brit. J. Educ. Psychol.*, 35, 1-8.

Ravenette, A. T. (1961). Vocabulary level and reading attainment: an empirical approach to the assessment of reading retardation. *Brit. J. Educ. Psychol.*, 31, 96-103.

Rutter, M., Tizard, J., and Whitmore, K. (Eds.) (1967). *Monograph on the Isle of Wight Study.* In preparation.

Schonell, F. J. (1950). *Diagnostic and Attainment Testing.* Edinburgh: Oliver and Boyd.

(*Manuscript received* 14*th July*, 1966)

BIRTH ORDER, ACHIEVEMENT-MOTIVATION AND ACADEMIC ATTAINMENT

By FRANK H. FARLEY
(*University of Wisconsin*)

Schachter (1963) has demonstrated that first-born persons exceed later-borns in academic attainment, and having rejected differential intelligence as an explanation, has suggested that motivational differences may account for the achievement discrepancy. An hypothesis that follows from Schachter's discussion is that first-borns will have higher need-achievement (*n* Ach) than later-borns. Sampson (1962) has reported results that do, and Wolkon and Levinger (1965) and Rosenfield (unpublished data cited by Wolkon and Levinger) have reported results that do not support a relationship between ordinal position and achievement motivation.

To provide further evidence on the motivational hypothesis, the following data were collected using forty-nine male university students and the twenty-eight-item forced-choice *n* Ach scale from the EPPS (Edwards, 1959). A major component of achievement motivation is fear of failure (Atkinson, 1958); to measure it in the present study, the test anxiety scale (TAS) of Mandler and Sarason (1952) was administered. A measure of academic attainment—Grade Point Average (GPA)—based in all cases on a four-year completed undergraduate programme in Education was likewise obtained.

Twenty-five subjects were first-borns, and twenty-four were later-borns; the mean *n* Ach score of the first-borns was 15.44 (SD=5·31) while that of the later-borns was 13·38 (S.D.=4·17). A *t* test yielded a ratio of 1·47 (48 df) which was clearly non-significant. This difference, though in the direction of higher *n* Ach in first over later-borns, as expected from Schachter's theoretical analysis, must not be taken seriously in light of the insignificant *t* value. The first- and later-borns did not differ significantly on such variables as age, family size and sex of siblings. TAS scores of the first-borns (mean=5·17) did not differ significantly from those of the later-borns (mean=6·56), though the direction of the difference favoured a lower fear of failure in first-borns, which would be expected from Atkinson's achievement motivation theory taken in conjunction with the hypothesis that first-borns are higher in *n* Ach than later-borns The GPA scores of the first-borns (mean=3·00) did not differ significantly from those of the later-borns (mean=2·94). The most specific conclusion here, within the limits set by sample size and composition, is that birth order and *relative academic achievement in university* are unrelated.

The results of this study are largely negative with respect to ordinal differences in achievement motivation and academic achievement as defined.

REFERENCES.

ATKINSON, J. W. (Ed.) (1958). *Motives in Fantasy, Action and Society.* Princeton Van Nostrand.

EDWARDS, A. L. (1959). *Personal Preference Schedule, Test Manual.* New York: Psychological Corp.

MANDLER, G., and SARASON, S. B. (1952). A study of anxiety and learning. *J. Abnorm. Soc. Psychol.*, 47, 166-173.

SAMPSON, E. E. (1962). Birth order, need achievement and conformity. *J. Abnorm. Soc. Psychol.*, 64, 155-159.

SCHACHTER, S. (1963). Birth order, eminence and higher education. *Amer. Sociol. Rev.*, 28, 757-768.

WOLKON, G. H., and LEVINGER, G. (1965). Birth order and need for achievement. *Psychol. Rep.*, 16, 73-74.

WASTAGE OF ABILITY AMONGST THE INTELLECTUALLY GIFTED

By H. G. ARMSTRONG

(*Educational Psychologist, West Riding*)

Summary. In a follow-up of children admitted to grammar schools in 1956 and 1957 with I.Qs. of 135+, more than twice as many from professional, as from working class, homes entered full-time higher education. Performance at Ordinary and Advanced levels of G.C.E., showed no significant differences between the two groups.

I.—Introduction.

In ' Early Leaving ' wastage amongst children of grammar school ability was closely scrutinised. The selective entry for 1946 was divided into a top, middle and lower third and progress assessed five years later according to a number of academic categories and parental occupation. The relevant table (Table K) shows that almost half (48·3 per cent.) of those who were in the lowest selection group at 11, but in the highest academic category at 16, came from professional homes, whereas more than half (54 per cent.) of those in the highest selection group at 11, but in the lowest academic category five years later were the children of unskilled workers. The Report observes: " It is beyond doubt true that a boy whose father is of professional or managerial standing is more likely to find his home circumstances favourable to the demands of grammar school work than one whose father is an unskilled or semi-skilled worker. The latter is handicapped."

In a follow-up of children born in 1940-41, the Robbins' Report (Appendix One, page 44) concludes : " Among the most able children, the differences between social classes were less marked at the stage of ' O ' and 'A' levels, but the proportion entering degree-level courses is twice as high for middle class as for working-class children." The same influences that affect progress in the grammar school are shown by Douglas to be at work in the primary school. Although it is not possible to apply the usual checks to Douglas's data, he rightly emphasizes the significance of parental attitudes and parental aspirations : " The middle class parents take more interest in their children's progress, the fathers, as well as the mothers, visit the schools more frequently, and they are more eager for their children to have a grammar school education. These differences are to be found within each social class and, when a re-classification is made according to parental attitude, those whose parents show more interest make greater progress."

II.—Academic record of gifted children.

In the 11+ age groups, for the years 1956 and 1957, there was a total of 849 children in the West Riding with I.Qs. over 135 as indicated in Table 1.

TABLE 1

Numbers of Children with I.Qs. of 135+.

Year	No. in Age Group	I.Qs. of 135	Moray House Tests	Proportion
1956	23,944	366	57,58	1·5
1957	24,602	483	59,60	2·0
Total ..	48,546	849	—	1·75

' O ' and 'A' level results are available for 577 or more than two-thirds of the total number given in Table 1.

In Table 2, the students have been arranged in two groups at the end of their grammar school course, according to whether they continued in full-time higher education or left school to take up jobs, and distributed according to parental occupation.

TABLE 2

CHOICE IN RELATION TO PARENTAL BACKGROUND. FOR (a) 1956, (b) 1957.

	Category of Parental Occupation*														
	I		II		III		IV		V		Unclassified		Total		Grand Total
Full-time	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)	(a) (b)
Higher Education...	21	14	59	58	82	107	18	18	1	5	11	4	192	206	398
Other Occupations	4	7	9	16	36	63	11	17	4	5	5	2	69	110	179
Total	25	21	68	74	118	170	29	35	5	10	16	6	261	316	577

* I—Higher Professional. III—Skilled.
II—Lower Professional. IV—Partly Skilled.
V—Unskilled.

Since all of the students listed under Table 2 may be regarded as intellectually gifted, it might reasonably be expected that a considerable number would go on to full-time higher education. It can be seen that 398 out of 577, or some 70 per cent., did so. The wastage of 30 per cent however, is somewhat misleading, as it conceals a differential wastage of some magnitude that exists between those from different backgrounds.

Taking the two years 1956 and 1957 together, the table shows that thirty-five out of forty-six, just over 75 per cent., of those from higher professional homes, go on to higher education compared with six out of fifteen, or 40 per cent., of those whose parents are engaged in unskilled occupations. This means that amongst the latter group, although all of then may be regarded as of university potential, the wastage is more than twice as great as that of the former. If the figures for I and II, are combined and contrasted, with those for IV and V the differential remains almost constant.

Table 3 below contrasts the performance of those who went on to full-time higher education, with those who took up other occupations.

TABLE 3

AVERAGE NUMBER OF PASSES AT 'O' LEVEL.

	No.	Mean	S.D.
Higher Education	398	7·38	1·74
Other Occupations	179	5·50	2·22

It is clear from the table that the better qualified group went on to higher education. The Critical Ratio is 12·86.

Table 4 below compares at the same level the professional (I and II) and manual (IV and V) groups :

TABLE 4

'O' LEVEL RESULTS OF PROFESSIONAL AND MANUAL GROUPS.

	No.	Mean	S.D.
Professional	188	6·98	1·82
Manual	79	6·46	2·52

As in Table 3 these results are for the years 1956 and 1957 combined. With a Critical Ratio of 2·60 there is no clear difference in performance between the two groups. The same holds true at 'A' level.

TABLE 5

AVERAGE NUMBER OF PASSES AT 'A' LEVEL OF PROFESSIONAL AND MANUAL GROUPS.

	No.	Mean	S.D.
Professional	163	2·80	1·56
Manual	50	2·82	1·02

With a Critical Ratio of 1·00 there is no difference in achievement between the two groups.

From the original number, 398, going on to full-time higher education, 259 or 65 per cent. take up university places. These are distributed in Table 6, according to parental occupation.

TABLE 6

NO. AND PROPORTION ADMITTED TO UNIVERSITY.

	Category of Parental Occupation						
Years	I	II	III	IV	V	U.C.	Total
1956, 1957	25	80	119	20	4	11	259
Proportions	9·7	30·9	45·9	7·7	1·5	4·2	100

It can be seen that 40 per cent. of the university places are taken up by the professional and just under 10 per cent. by the working class groups, although both may be regarded as of equal university potential.

III.—REFERENCES.

CURRY, R. L. (1962). The effect of socio-economic status on the scholastic achievement of sixth-grade children *Brit. J. Educ. Psychol.*, 32, 46-49.

DOUGLAS, J. W. B. (1964). *The Home and the School.* London: McGibbon and Kee.

Early Leaving (1954). H.M.S.O.

Robbins' Report (*Appendix One*) (1963). H.M.S.O.

COMMENTS

ON "AN INVESTIGATION OF THE CRAIG READER COURSE," BY K. D. DUNCAN AND R. FOSTER

By G. HARRY McLAUGHLIN

(The City University, London)

Duncan and Foster[3] stigmatise as 'disappointing' my finding[4] that undergraduates who completed the Craig Reader Course[2] failed to show a significant increase over a control group in the amount they could recall from a technical digest after five minutes reading. They contrast this with Poulton's survey[5] of forty reading improvement courses based on the Harvard films, in thirty-three of which there were general increases in comprehension, as measured by multiple-choice questions, though only three increases were significant ($P<0{\cdot}05$). Duncan and Foster go on to demonstrate that thirty-four soldiers who took the Craig Course scored significantly higher on Black's multiple-choice test of comprehension,[1] compared with a control group of fourteen.

However, it must be observed that the findings of Poulton and of Duncan and Foster in no way conflict with my own. They show that the training given by a reading course usually improves the ability of a subject to respond correctly to test questions on a passage *when the questions are so framed as to remind the reader of the content of the passage.* By contrast, my investigation concerned *unaided* recall, which is required in most reading situations in real life.

As one profesionally concerned with improving scientists'communication skills, I would be delighted if it could be shown that any reading course can improve unaided recall. Unhappily, my conclusion that the Craig Reader Course, at least, fails in this respect, is supported by Duncan and Foster's other results · there was no significant difference between the abilities of their experimental and control groups to answer *open-ended* questions on a pamphlet they had read for 15 minutes.

REFERENCES.

[1] BLACK, E. L. (1962). *Manual of Instructions for Comprehension Test for Training College Students.* London: National Foundation for Educational Research.

[2] CRAIG RESEARCH, INC. (1961). *Advanced Reading Program A.* Los Angeles: Craig. Corporation. Ashford, Middlesex: International Tutor Machines, Ltd.

[3] DUNCAN, K. D., and FOSTER, R. (1966). An investigation of the Craig Reader Course *Brit. J. Educ. Psychol.*, 36, 327-330

[4] McLAUGHLIN, G H (1964). Research Report *New Education*, 1, 1, 35-37.

[5] POULTON, E. C. (1961). British courses for adults on effective reading. *Brit. J. Educ. Psychol.*, 31, 128-137.

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES

PSYCHOLOGICAL FACTORS ASSOCIATED WITH THE SOCIOMETRIC STATUS OF CHILDREN ATTENDING A COMPREHENSIVE SCHOOL IN BRECONSHIRE

By HOWARD H. RICHARDS

(*Abstract of a thesis submitted for the degree of M.A. in Education of the University of Wales, at the University College of Swansea,* 1966)

AIMS.

To investigate the relationship of the sociometric status of 'A' stream children in a mid-Wales rural comprehensive school to their ability and achievement and their personality stability and to assess the influence of geographical factors on group formation.

THE SAMPLE.

Included 189 pupils aged 11 to 16, receiving a grammar type education at a county comprehensive school, 49 per cent. were boys and 51 per cent. girls, they came from a variety of socio-economic backgrounds, 57 per cent from rural and 43 per cent from urban areas

METHOD.

Sociometric status was derived from a McKinney[1] type sociometric test using reactions of acceptance, rejection and indifference to each classmate's becoming a member of a class work group. The sociometric status score was the algebraic sum of all the acceptances and rejections, and this was converted to a sociometric status index by dividing the raw score by N–1. A second sociometric test with three choices and two criteria (working and seating companions), was given six to eight weeks later and also used reasons for acceptance and rejection as employed in the McKinney two dimensional sociometric technique. The two tests showed a fairly high constancy of sociometric status over this period when the test results were intercorrelated.

Ability for each class was assessed by Raven's Progressive Matrices used as a group test while achievement scores were derived from School Christmas 1964 Examinations Stability of personality, or a measure of social adjustment in the classroom, was obtained from the Bristol Social Adjustment Guide—The Child in School (Stott and Sykes) completed by members of staff during the academic year 1964-65. Each endorsed trait on the Guide counted 1 point in the overall social adjustment score, where over 20 indicated maladjustment. Interests were assessed by a 130 item interest inventory prepared for the investigation and a width of interest score was derived. . Both sociographic maps and regional rural/urban sociograms were prepared in the analysis of geographical relationships.

ANALYSIS OF DATA.

The sociometric status scores and scores for all the variables were converted to rank orders and intercorrelated by rank order correlation, since sociometric status scores were not normally distributed. The analysis of cleavage into regional groups influenced by geographical factors was carried out by use of regional sociograms and *chi* square.

RESULTS.

(1) Sociometric status showed low positive correlations with intellectual ability (r=+0·21 to +0·63), but when broken down to groups of high, medium and low sociometric status, the high status groups always showed higher mean ability scores than the low group.

(2) Achievement scores were similarly related (r=+0·30 to +0·68) and again high sociometric groups always had higher mean ability scores than the low groups.

(3) Deteriorates in this study were generally in the lower half of the social acceptance scale (24 out of 37)

(4) Highest positive relationships existed between high sociometric status and high personality stability (r=+0·50 to +0·82).

(5) Maladjusted pupils (7·4 per cent. of sample) were on the whole an under-accepted group, the majority being in the lowest quartile of acceptance.

(6) A fairly common pattern of the unforthcoming, shy or ' recessive ' personality was evident in the maladjusted and unsettled pupils in this survey

(7) Interests did not appear to be positively related to sociometric status in this survey.

(8) Cleavage into town and country groups was evident in class sociograms and confirmed by *chi* square analysis. It was thought to be due to the cultural dichotomy between the anglicised town and the surrounding Welsh country districts.

CONCLUSIONS

(1) The relations of ability, achievement and personality stability to sociometric status are complex, but the study suggests that the most highly accepted pupils are generally those with the most stable personalities and are those most likely to be succeeding in their school work.

(2) The sociometric technique is a simple and useful diagnostic procedure for defining the ' outsider ' category from which come a disproportionate number of the maladjusted and potential, or actual, deteriorates in school

(3) The study emphasizes the value of appreciating the ecological factor in group formation Cleavage into regional cliques occur which may affect social integration within the classroom and even throughout the school.

REFERENCE

[1] McKINNEY, J. C. (1948). A two-dimensional sociometric test. *Sociometry*, XI, pp. 356-367.

BOOK REVIEWS

ANSTEY, E. (1966). *Psychological Tests.* London : Nelson, pp. xix+300, 63s.

This book is concerned with the details of test analysis for the test constructor who is mainly concerned with the application and the preparation of selection instruments for educational, vocational and industrial purposes It contains a wealth of practical experience gained over the years in the preparation of psychological test material at many levels for the Directorate for the Selection of Personnel (D.S.P.) War Office and for the Civil Service Commission Research Unit. The author deals in some detail with the writing of items, methods of item analysis (some developed in his own researches), job analysis, reliability and validity ; all this is most useful since practical applications of test construction should never be divorced from the underlying psychological and statistical theories.

In this account of the methods used for constructing psychological tests for selection purposes, only brief references are made to the need for norming devices ; there is a reference to the use of percentiles, but none at all to the usual methods available for obtaining standard scores and adjusting these scores for age allowances. For these the research worker will have to consult standard tests and the articles by Godfrey Thomson (*Brit. J. Educ. Psychol.*, 1932, 2) and D. N. Lawley (*Brit. J Psychol. Statistical Section*, 1950, Vol. III) Dr. Anstey's book does provide an up-to-date account of the scientific procedures for test construction and such a detailed account is long overdue, but there is also a warning implied by the author and emphasised by Sir Cyril Burt in his introduction—a test must be carefully constructed and validated to be of value. " The first duty of both the inventor and the user is to test the test." W. H. KING.

ATKINSON, J. W., and FEATHER, N. T. (1966). *A Theory of Achievement Motivation.* New York : John Wiley, pp. x+392, 87s.

This volume testifies to the significant advances which have been made in research and theories of achievement motivation since the initial studies set forth a valid means of assessing individual differences in need achievement (1953) ; these used the projective thematic apperception method (TAT) and demonstrated the effects of high and low motivation on performance. The editors, John W. Atkinson, one of the pioneer workers in the area, and Norman T. Feather, have brought together a lucid set of papers. They have both contributed considerably themselves to the development of the sophisticated, current thory and to the empirical evidence. Their theory which takes account of the tendency to avoid failure (T-f) as well as to achieve success (Ts) repays careful study. The repeated finding in studies of achievement and of level of aspiration that there are systematic individual differences in preference for intermediate risks *versus* erratic choice of extremes is neatly explained by their notions of subjective probability of success as a consequent of a particular activity (Ps) and incentive value of success in that activity (Is). Furthermore, they indicate experimental means by which these preferences can be manipulated. Thus, a link is formed between traditional descriptive assessments of individual differences and the control of independent variables usually associated only with experimental psychology.

Teachers should find papers of special interest in the section on applications and social implications. Chapter 15 deals with the vexed problem of streaming primary school classes. The theory of achievement motivation leads deductively to the hypothesis that students in whom the motive to achieve success is greater than the motive to avoid failiure should exhibit greater interest and learning in ability streamed classes. While their limited data from American sixth grade classes supported this hypothesis, the alternative that those in whom Maf is greater than Ms would lose interest and decline in achievement in homogenous groupings was only partially substantiated Loss of interest was not accompanied by performance decrement. Of practical concern too, is the relationship between the motive to avoid failure and unrealistic (in terms of ability) vocational choice reported in Chapter 12. The evidence indicates that achievement motivation as well as intelligence ought to be considered in making decisions on educational policy.

The editors' suggestion that the reader turn to the final synthesising chapter before tackling the entire set of papers is useful. The chapters have been arranged in chronological order and thus the earlier theoretical papers contain statements later amended in the wake of empirical evidence. Those interested in this area but who are unfamiliar with it may find the following texts helpful :

McClelland, D. C., Atkinson, J. W., Clark, R. A., and Lowell, E. L (1953). *The Achievement Motive.* New York : Appleton-Century-Crofts.

Atkinson, J. W. (1966). *An Introduction to Motivation.* Princeton . Van Nostrand.

Barbara B. Lloyd.

Berlin, I. N., and Szurek, S. A. (Edits.) (1965). *Learning and Its Disorders : Volume I—Clinical Approaches to Problems of Childhood.* California : Science and Behaviour Books, Inc., pp. 295+xviii, $5.95.

This is the first volume of *Learning and Its Disorders*, entitled *Clinical Approaches to Problems of Childhood.* However, as such, the title tends to be somewhat misleading, for, although its editors do adopt a clinical approach within the learning theory framework, the problems to which they refer are mainly of an anti-social, delinquent nature, and they concentrate practically exclusively on the adolescent stages, rather than the full age range of childhood.

The editors, two Californian psychiatrists, have presented this volume as a collection of published and hitherto unpublished papers. On this account the book has not at all times the continuity that might be hoped for. Complete entities unto themselves, the papers tend to read in a rather repetitive fashion when amassed in such book form, making the volume more of value as a reference manual, than as a standard text. Only in the latter part of the book do the papers seem to supplement and complement each other very well, as in the description of the San Francisco anti-delinquency school project.

As the introduction itself states, the papers deal in main " with the early infant-parent relationship requisite for optimal learning and healthy living ; the obstacles to learning, creativity and integrated living in society ; and the specific role of teachers in the vital learning experiences of children." Within this framework is emphasized the importance of realistic expectations on the part of the teacher, in his dealings with disturbed children, the necessity for him to express openly to his colleagues his understandable anger and frustration when failures occur, and the value of a ' firm fairness ' in class discipline. Berlin believes that authoritative, but non-authoritarian and non-retaliatory behaviour, will alter delinquent character structure. The authors also stress the importance of co-operation between teachers and mental health workers, so that the anxieties of administrators about teacher-parent-pupil relationships can be lessened. Therapy with the parents, especially the mother of a disturbed child, is advised, while it is suggested that the learning of academic material may be the most important therapeutic experience of all for the child himself.

These points make this volume sure of a place in the libraries of most educators, while the most important point of all, perhaps, is made in the description of the practical successful co-operation of school administrators, teachers, psychologists, social workers and psychiatrists in the therapy and prognosis of anti-social delinquent children.

Joan C. McQuoid.

BIDDLE, B. J., and THOMAS, E. J. (1966). *Role Theory: Concepts and Research.* New York: John Wiley, pp. xiv+453, 68s.

During the last twenty-five years, we have become familiar with the methods and theories of sociometry and group dynamics, and have learnt to regard human conduct as, in many cases, reciprocal action between individuals. Role theorists see these actions as organized into roles, and stress the importance of another kind of interaction, that between the role and the self. Role theory is a new field of study, and, if the editors are right in claiming that it is on the threshold of recognition as a specialized field, then this book is likely to be of considerable importance. The first four chapters, designed to articulate the domain of role theory and specially written by the editors, are followed by a selection of readings and an extensive bibliography running to forty-seven pages.

The first chapter is devoted to a discussion of the nature and history of role theory while the next three deal with its conceptual structure. Chapter II deals with concepts for classifying the phenomena of role, Chapters III and IV with those for the properties and variables of role phenomena. Grappling with an unfamiliar terminology is never easy, and these chapters are not easy reading. They are, however, rewarding, in that anyone studying them seriously may expect to gain a sound, systematic knowledge of the concepts with which they deal. These four chapters are relatively short, occupying only sixty-three pages in all, and the remainder of the book consists of extracts and abstracts from a wide variety of sources. The years of original publication range from 1937 to 1964, with only two predating 1949, and the selection may fairly be claimed to be as up-to-date as possible.

The readings are arranged in eight sections, each with a brief introduction. Earlier sections deal with simpler and more fundamental topics, such as positions, prescriptions, descriptions, and performance and interdependence. More complex and specialized topics, namely, differentiation, specialization, and division of labour, consensus and conflict, sanctioning and conformity, and learning and socialization, are treated later. It is unlikely that elementary courses will deal with role theory as such, though all students of educational psychology should be aware of the influence of the self-concept on behaviour. For those who wish to pursue the subject rigorously to a more advanced level, this book supplies both a text and a basic reference book, while the bibliography is invaluable as a relatively complete list of serious theoretical and empirical published contributions to role literature, though it deliberately omits such topics as role playing and psychodrama. K. M. EVANS.

CLYNE, MAX B. (1966). *Absent: School Refusal as an Expression of Disturbed Family Relationships.* London: Tavistock Publications, pp. 269, 39s.

Dr. Clyne writes from the point of view of a general practitioner who has treated a number of cases of school refusal for whom child guidance treatment was not readily available. Although the fifty-five cases which he examines in some detail were, on the whole, younger and less difficult than those normally seen at child guidance clinics, Dr. Clyne discusses the problem of school refusal very comprehensively. He has familiarised himself with a wide range of studies dealing with absence from school, and examines these with critical acumen.

Like many other writers on this subject, Dr. Clyne sees school refusal as the result of a disturbance of relationships within the family. While recognising that difficulties at school may occasionally be of importance, he does not consider that these are a major contributory factor in school refusal. In fact, the school situation was rarely investigated at first hand, and it seems a little strange that Dr. Clyne, after making some perceptive observations on relationships at school in the early part of the book, should not feel that the school situation merits the same kind of investigation as the home circumstances invariably received. As recent work (for example that by Terence Moore on primary school children) has emphasised, school can be a source of considerable anxiety to a number of children.

Dr. Clyne argues that the family doctor, with the co-operation of specialist medical colleagues, is the best person to treat school refusers. Since, however, he is clearly aware that emotional factors are not always taken into account when physical symptoms, such as those which are commonly presented by cases of school refusal, are dealt with, it is a pity that he does not discuss in more detail the implications of his argument for the training of general practitioners and for their collaboration with workers in other disciplines. He seems, too, to underestimate the role of the teacher and the educational psychologist in the readjustment of school refusers. Even in those cases where there are no initial difficulties at school, the child nearly always needs a skilled and sensitive approach when he returns to school, and the educational psychologist, surely, has an important contribution to make in advising on school placement and methods of handling the child at school.

Dr. Clyne himself shows rare clinical insight, and his book, although it is rather expensive, and contains little that is new, is well worth reading by all professionally concerned with children's problems. His methods of investigation may, as he himself states, have been somewhat haphazard, but the presentation of his views and findings is most systematic. His precision of thought and lucidity of expression, his reluctance to over-simplify issues and to make generalisations, as well as his skilful use of case material to illustrate his points all contributed to making the book a pleasure to read, even for a reviewer with a definite bias in favour of the team approach of the child guidance service. MAURICE CHAZAN.

CONGER, J. J., and MILLER, W. C. (1966). *Personality, Social Class and Delinquency*. New York: John C. Wiley and Sons, pp. viii+249, 60s.

This book is a report of research carried out on youngsters in the city of Denver, Colorado, U.S.A., with 220 pages, eighty-two references and three appendices. The main hypothesis tested in this study was that there are interactive relationships between Intelligence, Socio-economic class background (here called ' deprivation '), and Personality and Delinquency.

The researches started with a representative sample of Denver school children in the ninth grade of Junior High School and subjected them to a lengthy battery of personality tests. Information was then obtained from courts about who, from this sample, had " been accepted by the Juvenile Court." This, as a criterion of delinquency, was rather less than the total number reported for delinquency and rather more than those convicted by a court One-hundred-and-eighty-four members of this sample of 271 delinquents were then matched with a non-delinquent peer of the same I.Q., socio-economic class, age, ethnic origin, school attended, and area of home—an impressive matching procedure, in which it had been impossible to find a match for eighty-seven boys who were then discarded from the study, the analysis being carried out on the remaining 184. No information is given about the effect of this selection from the delinquent population on the distribution of the variables under consideration.

The statistical model used was a 2 × 2 × 3 analyses of variance design with two levels of delinquency (delinquent *versus* non-delinquent), two levels of socio-economic status (deprived *versus* non-deprived), and three levels of I.Q. (high, medium and low), with personality ratings and test scores as the dependent variable. This design was attenuated in many ways, through lack of appropriate subjects to fill particular cells; one of these truncated designs included a 2 × 3 analysis (delinquency as one dependent variable and a combination of I.Q. and socio-economic status as the other)—the authors recognise impropriety of such a compounding of variables, but still proceed to interpret the results No use was made of analysis of covariance or discriminant function

Two findings stand out as of importance · first, there was overwhelming evidence, from very careful analyses of ratings made by teachers at various stages of the boys' school life, that the delinquents not only were less well adjusted to school, less responsible and friendly to their peers, more aggressive and resentful to authority, but also, they always had been so at school, even though their delinquencies did not become serious enough to be reported until they were between 12 and 14 years of age.

Secondly, the results of all the interaction-analyses suggested the very powerful influence of intelligence, rather than either social class or personality.

It is stimulating but unsatisfying to the professional researcher. Too little information was provided—no data of the analyses of variance or factor analyses were provided—only conclusions. We were left to assume that homogeneity of variance was tested for, and that a sensible criterion for number of factors to be extracted was used. Similarly, not all the data make sense, for instance, one factor analysis accounted for 128 per cent. of the common variance, and even allowing for the obvious misprint on factor 6 (31 printed instead of 13), they still managed to explain 110 per cent. of the variance with orthogonal factors ! The general reader will probably find the excessive reliance of diagrams and the analysis of variance format disconcerting, and in such a relatively short book there is not room for compensating discussion of the issues raised by the findings, especially the one showing that nursery school teachers' ratings of behaviour can be powerful predictors of delinquency occurring as much as six or eight years later. A. P. SEALEY.

DELACATO, CARL H. (1966). *Neurological Organization and Reading.* Illinois: Charles C. Tromans, pp. x+189, $7.50.

In two earlier publications (in 1959 and 1963) Delacato put forward his view that language problems can be diagnosed, treated and even prevented through the modification of neurological organisation. In this present book he outlines his theories, declaring them in evolutionary terms. His views may be summed up in the following quotation : " The basic difference between man and the animal world is that man had achieved cortical dominance wherein one side of the cortex controls the skills in which man out-distances lower forms of animals. This whole phylogenic process is recapitulated ontogenetically with each human being. In the event that there is some obstruction to this ontogenetic recapitulation, communication and language dysfunctions occur " (page 7).

In his claim that difficulties of speech and reading may be different manifestations of the same problem, Delacato would be supported by some ; while his view that doubtful laterality may be in some way connected with this is not unsupported by neurological evidence. He goes further, however, and claims that the ideal is that dominant hand, foot and eye be on the same side—and if this is not found to be the case, then steps should be taken as a remedial measure to train the use of the ' proper ' hand and to make the eye and hand on that side also dominant. His diagnostic tests also include analysis of cross-pattern walking, crawling and of sleeping position—should these not measure up to the required patterns, remedial work is also recommended in these ' lower ' areas of neurological functioning.

Part II of the book is described as a ' philosophic and logical overview,' and presents at greater length his views on the evolutionary development of speech. He claims to have found 60 per cent. of good readers completely one-sided and only 6 per cent. of poor readers, indicating the place of neurological dysfunction in reading disability.

In the remainder of the book ten experimental studies edited by Delacato are presented, lending support to the theoretical framework he has outlined.

This book is not in any sense a textbook on reading. Much of what is said, though stated dogmatically, is mere speculation. It might perhaps be of interest to anyone particularly concerned to keep abreast with new theories in the field of reading ; but in the hands of anyone less well versed in the field it would do more harm than good. The sceptic is in no way reassured by the studies claimed to present a scientific appraisal of the theory. In editing them, Delacato states that he omitted some of the biographical material but left them otherwise untouched. There still remains a considerable amount of repetition, yet the studies on the whole lack clarity. One is left with the feeling that the improvements (slight in some cases) which the experimental groups gained over the control groups after a period of training to improve dominance and thus neurological functioning, could equally well have been the result of any of a number of factors not lending support to Delacato's theories. MARGARET M. CLARK.

FERNALD, L. D., and FERNALD, P. S. (1966). *Overview of General Psychology.* New York: Houghton Mifflin, ; London: Harrap, pp. 354, 21s.

Particularly in America, the size of introductory textbooks in Psychology has tended to increase as more and more detail is added to the central core. The authors of this book claim that while this may eventually lead to greater understanding, in the early stages, the student may become confused and overwhelmed by the mass of facts which may appear to him to be unrelated. In these circumstances he may fail to detect the basic general principles which might have been meaningful to him had they not been buried under the mass of supporting researches.

To counter this, Fernald and Fernald have produced this ' Overview ' which is, in fact, a short linear programmed text designed to cover the major basic principles taught in the fifth edition of Munn's *Psychology* (Munn appears to have given his blessing to the venture). As such, it is broken into eighteen sections which follow the chapter headings of Munn and a careful study of five of these with a group of mature students showed that the coverage of the significant principles was quite comprehensive In an appendix, the average time taken by American first-year undergraduates to complete each chapter is given as a little over half-an-hour and a limited try out in this country with teachers who claimed to have read Munn as a preliminary to starting a course of psychology took about the same time. The error rate is very low; the authors found 5·2 per cent., we found it rather lower.

Long experience of teaching psychology makes this reviewer agree with the authors that students often do become confused and fail to see the underlying principles for the welter of explanatory detail. Hence, he welcomes this book as a means of helping students to develop their basic psychological concepts and clarify their minds about their essential relationships. With English students it will be most valuable, perhaps, as a means of revision at the end of the first year of a course in psychology, and though it deals with general psychology almost all of it is relevant, for students of educational psychology. The principles taught in it are so general that students who have used other books than Munn's will be at no serious disadvantage and there is, in an appendix, a chart showing the cross réferences of chapters in Munn and nearly thirty other texts.

Not really a library book, but one for personal possession to be marked, annotated and cross referenced by first-year students. L B. BIRCH.

ILLINGWORTH, R. S., and ILLINGWORTH, C. M. (1966). *Lesons from Childhood.* Edinburgh: E. and S. Livingstone, pp. xi+384, 37s. 6d.

Your reviewer came to this book with the highest expectations. The title and conception of the work is challenging and not otiose. Its authorship by two leading paediatricians, paricularly interested in the developmental approach, augers well. Hopes are inevitably entertained that the volume will prove, in its different way, as suggestive and important as Gesell's *Biographies of Child Development* or Terman's work on the gifted child. What, in fact, do we find?

The bulk of *Lessons from Childhood* is devoted to describing the early life of people who later achieved distinction in a variety of spheres. No less than 450 biographies and autobigraphies are ransacked for data, which are arranged under various heads in fourteen chapters. The stories and anecdotes have a gossipy fascination and many less well known facts are presented. A vivid impression of the great unhappiness and wanton mismanagement of the childhoods of many of the ' unusual '—famous and infamous—subjects is left with the reader. Looking for the ' lessons,' these take the form of tentative discussions and comments at the close of each chapter, and in a final passage of ' conclusions.'

Many fundamental questions of the greatest interest are raised, but the evidence as presented can only pose, not answer, these The auxiliary methods of scientific enquiry are scarcely employed. There is for instance, no definition of the term ' unusual ' and the celebrity of those included has a vast and almost ludicrous range. The anecdotal material is left to speak entirely for itself and comparative information from social history or other sources is not used. As a result of all this, the final conclusions are unproved, rather than certainly or probably true or false. They are that there is an ' absence of pattern ' in the biographies and that it cannot be said whether the people described achieved fame in spite of or because of their circumstances. In view of this, the book's ultimate 'lesson' counselling ' tolerance,' is hardly scientific, though it is certainly humane.

Summing up, the book can be described as worth while in intention, but unsatisfying in outcome. Its objectives are significant and original, but the information resulting from the authors' labours is baffling rather than enlightening. We are still without an answer to that perennial question concerning the origins of greatness: what porridge had John Keats? O. C. SAMPSON.

JORDAN, THOMAS E. (1966). *Perspectives in Mental Retardation.* Illinois: Southern Illinois University Press, pp. xiii+358, $8.0.

This is an exceptionally wide ranging anthology. It opens with an extract from John F. Kennedy's " National Program to Combat Mental Retardation," with its forceful plea for more research, more expenditure, more community care and a diminution of institutional treatment wherever possible The book then advances to some thirty or so articles or extracts, mainly culled from American sources, but not excluding contributions from this country, covering discussions of retardation in the context of the legal system, the family, institutional life, and the world of work. The treatment is interdisciplinary and contributions are to be found from many branches of medicine and the social sciences. Content is divided into eleven sections, each with three articles. Thus, Section 5, ' Language,' consists of a case study of speech therapy techniques used with an adult imbecile, a survey of rates of progress in linguistic development in an institutionalized population of mongoloid children, and a neurological study of a case of congenital aphasia. Although references for further reading are appended by the editor, it is clear that the articles in any one section cannot provide a balanced coverage in another section. For example, Kirk and McCarthy's ' Illinois Test of Psycholinguistic Abilities " is included in the section on Instruction rather than in the language section. However, the professed aim of the collection is not depth, but width of treatment, partly to illustrate the extent of the field as a whole and partly to draw attention to the need for developing a methodology that is truly interdisciplinary. On this latter point, the editor dilates both in the preface, and, more extensively, in the contribution which he himself makes on " Constructs and Data Language in Special Education." The danger, he argues, is that in attempting to think about special education in terms of a variety of constructs culled from separate disciplines, each with its own relevance within that discipline, we may finish up " adding inches to ounces."

What the reader is offered, then, is variety of content and treatment so that he may both broaden his horizons and feel challenged to search for what is really relevant if research in mental retardation is to be related to the problems of special education. " Meaningfulness becomes the touchstone of serious scholarship: it functions much as Occam's razor in reducing the number of constructs which are extant at any given moment." Happily, the razor, though it has pruned, has not excised Seguin's eloquent endorsement of the rich communication-potential of ' restricted code ' language. With perhaps less justification, the razor has missed Langdon Down's magnificently absurd ethnic classification of idiots.

RONALD MORRIS.

LUNZER, E. A. (1966). *The Manchester Scales of Social Adaptation.* National Foundation for Educational Research.

Hitherto, the only comprehensive scale available for the measurement of social competence, has been Doll's Vineland Social Maturity Scale; the first version of which was published in 1935. This scale has proved of considerable value, especially in research and in the assessment of the needs of handicapped and deprived children.

It has, however, a number of defects, including inadequate standardisation, the small number of items alloted to each age group (particularly in the case of children of school age), the vagueness and subjectivity of many of the scoring procedures, and, for users in this country, the fact that its norms are based on an American sample. A new scale of social maturity has, therefore, been very much needed here.

Dr. Lunzer, fully acknowledging his debt to the Vineland Scale, has set out to provide a revised scale, based on an adequate sample of British children. His version concentrates on children of school age and covers the age range of 6 to 15 years. Containing eighty-eight items altogether, the scale consists of two separate parts, one assessing ' Social Perspective,' the other ' Self Direction.' Though no fundamental dichotomy between the two parts is intended, the first is made up of five sub-scales bearing particularly on cognitive aspects of social competence (e.g., general knowledge) and the second, also of five sub-scales, is concerned with the more practical aspects (e.g., social skills and activities). Most of the test items take the form of questions to be answered by the children themselves in a standardised interview. Raw scores can be converted into percentile norms for each of the ten sub-scales, for social perspective and self-direction separately, and also for the overall total. Apart from some misprints, the manual is a good one, being clear and informative. In the case of most items, we are told at which age 50 per cent. of a representative sample may be expected to pass The standardisation procedures are quite fully described, and there is an interesting discussion of the concepts of social maturity and social competence.

Dr. Lunzer's new scales are very welcome and he is to be congratulated on the progress he has made in a very difficult field. Most of the test items are easy to administer and in a form likely to produce a response from all but the most inhibited subjects. The availability of norms for each sub-scale and for each of the two parts is extremely valuable in that a profile of the child's social abilities is provided.

While detailed comments should be withheld until the scales have been tried out over a wider area, some criticisms may be tentatively proffered at this stage. The standardisation sample (of twenty boys and twenty girls for each year), is hardly adequate. The scoring, while aiming at objectivity combined with some flexibility, is still somewhat subjective and complicated in places ; and success in many of the Social Direction tests is dependent on what the child says he can do rather than on evidence of actual performance of social skills. The criteria for the cessation of testing in several sub-scales seem unduly harsh, seeing that a wide range of interests is covered and that the questions are not always in strict order of difficulty. Some of the items, too, in the Social Perspective section appear to test specialized interests, and their place in a general scale of social adaptation may be questioned.

However, Dr. Lunzer makes it clear that he regards his present scales as provisional and that he intends to revise them after continued work. They should prove a valuable diagnostic instrument, and now that the scales have been published by the National Foundation for Educational Research, we look forward to further reports on their use.

MAURICE CHAZAN.

MONTMOLLIN, ED. G. DE (1900) *International Journal of Psychology*. Paris: Dunod, Vol. 1, Nos. 1 and 2, pp. 1-86 : 87-176.

New journals in psychology have been springing up in all directions. But the *International Journal of Psychology* has the distinction of being sponsored by the International Union of Psychological Science, with the aim of being " devoted to cross-cultural comparative and co-operative research in general, genetic and social psychology throughout the world."

The first two numbers of the first volume have certainly attracted a world wide authorship, including Piaget, Jakobovits, Nuttin, Cesa-Binachi, Drever, Greenfield and Bruner, Frijda and Jahoda. They have written on such issues as " methods of cross-cultural research," " Japanese and American values," and " culture and cognitive growth."

The International Union of Psychologists has had an eventful fifteen years' history, which is briefly related by Roger Russell. Its latest achievement was the Moscow Conference in 1966 and it will be organizing the next international conference in London in 1969. It is to be congratulated on the successful launching of its journal and on its appeal, addressed to the thirty national members' societies, "that they, as psychologists, actively direct their professional energies toward reducing the menace of a new war, and toward improving the chances for peace throughout the world.

Two varieties of action are suggested The first is to emphasize certain conclusions reached by the overwhelming majority of psychologists but not always generally known or accepted These would include, for example, the conviction that there are no immutable tendencies or instincts in man which impel him inevitably toward war and destruction; that there is no fundamental inferiority of certain races, nations or social groups which makes them incapable of economic and social development, that there are no immutable human characteristics which render impossible co-operation and understanding between different races and nations.

Secondly, there is the need for research, and still more research What psychologists know represents only a fraction of what is needed It is, therefore, urged that psychologists conduct and facilitate scientific study of the attitudes and prejudices which endanger peace, of the psychological factors which interfere with co-operation and understanding, and of the techniques and experience which facilitate such understanding." This is an appeal to which we can surely respond

H. KAY

NISBET, J. D., and ENTWISTLE, N. J. (1966). *The Age of Transfer to Secondary Education.* London: University of London Press, published for the Scottish Council for Research in Education, pp. 5+199, 12s. 6d.

Educational psychologists are frequently asked to pronounce upon the question of whether there is any psychological justification for transferring children from primary to secondary schools at the age of 11 plus. In order to answer it they have have to collate evidence from the literature on growth studies, cognitive development and comparative education, but here at last we have a book which presents under one cover a scholarly survey and assessment of the evidence.

The research upon which Professor Nisbet and Mr Entwistle are engaged is based upon a follow-up of a year-group of Scottish children from primary to secondary school. It is still in progress and this is an interim report. The basic question posed is · "Is there a particular stage in child development when a change in education is appropriate—a time for a new approach, an increase in pace, the introduction of new subjects or of specialist teachers?"

The stage most commonly identified is puberty, and Chapter two examines the course of physical and mental growth at puberty concentrating in particular upon the evidence for a mental growth spurt. Examination of individual growth curves, sex differences and differences in test scores between early and late maturers leads to the conclusion that if any spurt exists, it is small and educationally insignificant. Chapter three examines the idea of adolescence as a unique period of psychological development and neatly brings together the Hadow recommendations and Stanley Hall's concept of adolescence as a 'new birth' These are discussed in the light of modern knowledge about social and emotional maturity, and Piaget's work. Chapter four describes the historical development of the 'clean break' system and shows how transfer arrangements have been the result of administrative considerations which no longer necessarily apply. Chapter five puts the problem in perspective in relation to practices in other countries and the final chapter looks afresh at the meaning of primary and secondary education.

The authors conclude that the existence of wide individual differences in maturation and the continuous nature of development make it advisable that the transition stage should extend over the whole period from age 10 to 13 By whatever method the schools are organised, the morals to be drawn have to do with more flexibility in transfer arrangements and more consultation between teachers at different stages

This book is an exceedingly clear presentation of a complex problem which is currently of great interest in view of the Plowden Committee's report It can be recommended as a real contribution to an area in which facts are difficult to disentangle and speculation is rife It should also be regarded as recommended reading for all teacher training courses on adolescence. P. K POPPLETON.

PEREZ, JOSEPH, F. (1965). *Counselling: Theory and Practice.* U.S.A : Addison-Wesley, Mass., pp. v+186, 38s. hardbound, 23s. softbound

This book is intended by the author to provide basic useful reading for those who counsel or intend to counsel others, be they guidance counsellors, social workers psychologists or psychiatrists Psychotherapy and counselling are seen as indistinguishable.

The author chooses to present his material within a frame of reference which stresses the importance in counselling of coming to terms with the client's basic psychological needs and of understanding the defences he uses both to defend and to satisfy these needs Following Maslow, four needs are defined, protection, love, self-esteem and self-fulfillment ; defences are divided into three categories reality-distorting (e.g , projection), rechannelizing (e.g., displacement) and reconstituting (avoidance and withdrawal). The individual's problems are a function of his actual or felt inability to utilize defences effectively when dealing with reality, and the goal of counselling is seen as helping the counsellee to learn to deal more effectively with reality The counsellor must understand how an individual's defence system is influenced by his perception of reality, his mode of interaction with reality and the meanings which he derives from this interaction , he must also come to understand how these elements function in his own personality. Desirable counsellor qualities of spontaneity, warmth, tolerance, respect and sincerity facilitate the interactive process between the client in need of assistance and the counsellor properly educated to give this assistance.

All these issues are discussed in some detail, interlarded with brief quotes from a wide range of recent research writings and illustrated with examples from interview sessions. The topics of communication, transference and insight receive special attention. In Part Two, three transcripts of counselling interviews are presented, each of which is analysed within the framework described in the main part of the book.

This book is an ambitious attempt to produce a text acceptable and useful to everyone in psychotherapy and counselling This is too great a task for any author. Both the level of discussion and the attempt to link up with research are quite often of superficial quality and at times Perez seems torn between eclecticism and a need to elaborate his own particular approach. On the other hand, the book is lucid, well-organized, contains well-written expositions of basic topics in the field and provides a useful framework, within which to view the complexities of any client /counsellor relationship

It is particularly difficult to evaluate for a British audience as the backgronud against which a book of this kind must be judged does not yet exist here. It is well-worth consideration, however, in a rapidly altering climate of opinion where counselling is in increasing demand and where it is being recognized, however grudgingly that no counselling other than the mere giving of information or advice can limit itself to the conscious layers of personality. P. HUGHES.

SARASON, I. G. (1966). *Personality: An Objective Approach.* New York: John Wiley, pp. xvi+670, 65s.

According to the author, this book on personality, aimed at American undergraduates, does not seek to be encyclopaedic, yet its section headings alone—theories of personality, assessment of personality, experimental study of personality,

personality development and, finally, deviant behaviour, indicate that it covers a very wide field. Indeed, it is, perhaps, this very breadth that is its weakness for each of these sections deserves a volume to itself and in reading them inevitable, if unfair, comparisons continually occur with such more detailed books. The first two sections indicate this very clearly.

The first section on theories of personality embraces in 100 pages Freud, Jung, Adler, the Neo-Freudians, learning theorists, Allport, Gestalt psychologists and field theorists. Such a task might well have taxed Procrustes and it would be difficult to refer students to this section in preference to larger works. Again, in the chapter on assessment only well known tests such as the M M.P.I , the C.P.I., the Rorschach and the T.A.T. are discussed which means that the ground is by now well trodden. Those of a factor analytic persuasion may also be surprised by the fact that the work of Cattell and Guilford merits exactly the same space as that of Ferguson (one paragraph). This section also exemplifies a lack of precision of style which, it might have been thought, would have been essential in a work of this kind. Thus, in discussing the M.M.P.I there is constant reference to 550 items when this figure, besides being wrong, is quite irrelevant.

On the credit side, as is usally the case with these books, there is an extensive bibliography which includes many recent experimental studies so that the book would be of value to specialists in different fields who wanted a reference book to modern work on personality. It is also refreshing to find a section on psychiatric personality disorders, neuroses and psychoses in a primarily psychological text. Although, by reason of length, it cannot compare with the books devoted to this subject it is an adequate outline which gains considerably from being linked to the psychological topics already discussed Too often is psychiatry entirely divorced from psychology.

In summary, it would be difficult to recommend this book to students, for the reasons of length indicated above. It falls, in fact, between two stools, being too long and complex as an introduction to the subject and too diffuse for detailed studies. Yet to survey so wide a field in such a span is no mean achievement and it is a pity that, in Britain, at least, it serves so little purposes PAUL KLINE.

SCHRAMM, W., LYLE, J., and PARKER, E. B. (1966). *Television in the Lives of our Children.* California: Stanford University Press, pp. iv+324, 22s. 6d

This useful paperback is a reprint of the authors' survey of 2,000 parents and 6,000 children in North America, which was first published in 1961. It is written in a lively popular style and gives a well-ordered, readable account of a great deal of information collected in eleven very different studies. As in Himmelweit's Nuffield study, information was collected by interview, questionnaire and personal diary though Schramm used rating-scales and vocabulary tests as well He was able to show, for example, that the early superiority in vocabulary of young children with television at home was lost as age advanced He recorded carefully the effects of age, ability, cultural setting, social-class, and aspirations on the 'patterns of television usage,' but did not undertake as much 'cross-cultural' comparison of children with and without television as did Himmelweit. Statistical material has been neatly condensed into a hundred page appendix which is a model of its kind.

There is a thoughtful and hard-hitting section addressed 'to broadcasters, to parents, to schools and to governments,' in the hope that they will mend their ways, and a useful annotated bibliography of ninety-one other television studies up to 1960.

Schramm's conclusions on children's experience of television are reassuringly similar to those of Himmelweit. Thus, Westerns do not frighten—nor do guns. It is daggers—and danger to animals—which cause distress. Again, girls get worried about married life which so often seems distressing on the screen. The short section on the 'behavioural effects' of television is less convincing—gathering together a number of opinions from a variety of sources rather than presenting research results. Schramm is quick to point out, of course, that we must not look on television's ' effects ' as uniform in all children—that a child ' uses ' television and selects ideas very different from those chosen by his neighbour. But, as Maccoby has said, this does not excuse us the important task of studying general effects on attitudes and behaviour. A great deal of the book is highly sophisticated ' audience-research relating the characteristics of viewers to the characteristics of the programmes they watch, and not enough is devoted to effects on attitudes and behaviour—but, perhaps, it is unfair to expect this from survey research ?

Alongside Himmelweit's book and Halloran's " Working Papers," this book by Schramm is required reading for all who would understand the impact of television on children. British readers should not be deterred by its American subject-matter, its early date or its popular style. JOHN LIGGETT.

WATTENBERG, W. W. (1966). *Social Deviancy Among Youth.* 65th Yearbook of the National Society for the Study of Education. U.S.A.: University of Chicago Press, pp. xi+434, +vi, 41s.

Much has happened since the National Society produced a Year Book on Delinquency in 1948, and it has become increasingly obvious that the term ' juvenile delinquency ' is an easy generalisation rather than a useful concept. Hence, this entirely new book on Social Deviancy, which deals with delinquency as one of a cluster of related problems, all connected with the increasing alienation of Western adolescents from mainstream adult culture. The book has fifteen chapters, each with different authors, but the usual repetitiveness and loss of impact does not appear, partly because the writers seem agreed as to the basic dilemma of urban American youth, and partly because the chapter themes are clearly differentiated and follow a systemic plan, beginning with reviews of recent research and development and leading to detailed examinations of what exactly schools must do to provide a compensatory education for the socially deprived and deviant.

By concentrating on the various forms of deviance that are socially undesirable, but not the result of major mental disturbance, the writers are able to include a variety of behaviour problems concerning teachers and employers as well as the police. The recurring theme is that in a society where young people are needed rather as consumers than as producers or responsible family members, it is increasingly difficult for them to achieve adequate ego development through a work role or social role. Whereas most middle class children (coarse classifications by class are over-used by nearly every contributor), largely conform to a tradition of deferred status, working class adolescent boys, with no such tradition and no compensatory gratifications like cars and generous allowances, are far more likely to opt out, either as anti-social deviants turning to aggression and deliquency as alternative sources of status in delinquent sub-cultures, or as a-social deviants, less alarming but eventually dangerous, whose solution to their existential problem is to retreat from social responsiblity into a Beat-type privatism.

The prediction, prevention and cure of both the above types of deviancy are the themes of the Year Book, the first chapters of which deal with recent theoretical changes in the approach to deviance and evaluate the resultant field work, impressive in variety and scale and depressing in the paucity of hard positive results.

Interestingly, both sociological and psychological contributors agree that the ordinary day school provides the best place for attacking deviance and that because of such recent changes in individual therapy as the concentration on ego functionining and ' coping techniques ' rather than on the id, on behaviour therapy rather than on depth analysis, the teacher, as the promoter of learning, will tend to be the key figure in such an attack. The second half of the book is largely concerned with how schools must re-organise themselves to cope with this, and there are a series of chapters on such topics as the value of specialist classes for disturbed children, various types of work experience courses, what exactly should be provided in compensatory pre-school classes, the value of ability grouping (!), what and how you teach to promote social commitment and so on. The flexible and imaginative efforts that have resulted from a non-ideology bound approach are most impressive, as is the last chapter on how the training and re-training of teachers should be affected by this wider concept of the role of the school.

Like most Year Books, this one covers its field exhaustively, but does so in a more stimulating and practical manner than many of its predecessors. It can be read profitably by any one interested in delinquency or problem behaviour.

RALPH D. DUTCH.

YATES, ALFRED (1966). *Grouping in Education*. New York: John Wiley and Sons, Stockholm: Alquist and Wiksell, pp. 314, 96s.

This important book, sponsored by the UNESCO Institute for Education, Hamburg, argues that all systems of mass education face a common problem—to find a basis for grouping pupils together, both in schools and in sub-groups within schools.

Part I consists of an analysis of delegate contributions to a conference held in Hamburg (1964) on national systems of educational grouping. After an account of principles in use, research data regarding the effects of the various practices are summarised. The conclusion is reached that apart from geography and age, no grouping criterion has been found to work consistently to the advantage of pupils and that some criteria are harmful.

There follows an interesting discussion of the factors affecting grouping arrangements and of their relation to educational ends and means. Investigation of sociological variables is recommended but work already done (e.g., R. H. Turner's analysis of British and American systems) gets little mention and this weakens the discussion. Thus, we learn that "intelligence is no longer regarded, as it used to be, as a fixed, unitary and hereditarily determined quantity" (p. 86); undoubtedly this underestimates the resistance of educational ideologies to new knowledge. By contrast, the treatment of teachers as a conservative interest group is illuminating and the suggestion that both large and small-scale grouping should be influenced by the findings of 'built-in' research rather than *a priori* 'principle' offers the basis for a breakthrough in the organisation of education. "... it is not grouping *per se* that produces educational results; it is rather that a certain composition of the group facilitates certain kinds of educational process" (p. 104). In turn, these processes can serve the alternative ends of social stability of social change and of individual uniformity or diversity.

In Part II specialist contributors elaborate aspects of the general theme, including Vernon who gives a concise summary of this position and Pasow, who unravells "The Maze of Research on Ability Grouping." In addition, there are 125 pages of abstracts which, with other references, bring important and unfamiliar European research to the attention of a wider public.

The compilation of conference papers into a comparative study of educational systems inevitably leads to omissions (France and Russia are hardly mentioned). Moreover, the attempt to impose a patterned interpretation upon a mass of findings gives the false impression that some contentious issues have been resolved. As a stimulus to new thinking in educational research and policy-making, however, the book is most welcome.

D. J. LEE.

NEW JOURNAL

NEW ZEALAND JOURNAL OF EDUCATIONAL STUDIES

Many readers to this *Journal* will be interested to see that there is a new journal being published in New Zealand under the Editorship of Professor P. J. Lawrence of the University of Canterbury. This journal has been established for the purpose of providing an avenue of publication for educational scholarship, research and critical comment and has special reference to the New Zealand and

Pacific area. It is not confined to one particular aspect of education, but welcomes articles in the field of educational psychology, sociology or methodology and in child and adolescent development. It also deals with administration and organisation, and historical and comparative studies in New Zeland and the Pacific area. Papers for publication in any of these fields should be sent to the Editor, The Education Department, The University of Canterbury, Christchurch, New Zealand. The first number, which came out last year, contains, first of all, a symposium on the scholastic achievement of Maori children. There is, first of all, a Historical Review of Policies and Provisions, by J. M. Barrington, The Scholastic Achievement of European and Maori Children, by M. N Lovegrove and then an article reporting the application of Reading Tests with Maori Children.

Among the other papers, one by P. S. Freyberg on Stages in Cognitive Development—General or Specific, will be of considerable interest to readers of this *Journal*. In it, Freyberg investigates the hypothesis that Piaget's stages are not as general and distinctive a group of intellectual phenomena as he has frequently claimed. The results of a longitudinal study carried out over a substantial group of children for over two years would seem to indicate that the picture of cognitive development is one of continuous growth rather than of discreet stages.

The new journal is to be issued twice yearly, in May and November, and the subscription rate is N.Z. 30s. per year, post free, payable in advance. Subscriptions should be sent to the Buisness Manager, P.O. Box 1540, Wellington, New Zealand.

L.B.B.

PUBLICATIONS RECEIVED

P. Seed (1966). *The Psychological Problem of Disarmament.* London : Housemans, pp. 74, 5s.

Thomas E. Jordan (1966). *Perspectives in Mental Retardation.* Illinois : Southern Illinois University Press ; U.K. Agent : Transatlantic Book Service, pp. xiii+358, $8.00.

W. H. Ehlers (1966). *Mothers of Retarded Children.* U.S.A. : Charles C. Thomas, pp. xxi+166, $6.75.

Jonas F. Soltis (1966). *Seeing, Knowing and Believing.* London : George Allen and Unwin, pp. 156, 25s.

Carl. H. Delacato (1966). *Neurological Organization and Reading.* U.S.A. : Charles C. Thomas, pp. x+189, $7.50.

Paul Nash (1966). *Authority and Freedom in Education.* New York : John Wiley, pp.x+342, 53s., 30s.

J. J. Conger (1966). *Personality, Social Class and Delinquency.* New York : John Wiley, pp. viii+249, 60s.

Max B. Clyne (1966). *Absent.* London : Tavistock Publications, pp. xi+269, 38s.

A. Yates (1966). *Grouping in Education.* New York : John Wiley, pp. 314, 96s.

Grant Fairbanks (1966). *Experimental Phonetics : Selected Articles.* U.S.A University of Illinois Press, pp. xi+274, 49s

Robert B. Textor (1966). *Cultural Frontiers of the Peace Corps.* U.S.A. : The M.I.T Press, Mass., pp. xix+363, 53s.

C. W. Taylor (1966) *Instructional Media and Creativity.* New York : John Wiley, pp. xxii+397, 68s., 30s.

Edwin E. Ghiselli (1966). *The Validity of Occupational Aptitude Tests.* New York . John Wiley, pp. vii+155, 60s.

Morris, I. Stein (1966). *Volunteers for Peace.* New York : John Wiley, pp. xiv+258, 60s.

G. O. M. Leith (1966). *A Handbook of Programmed Learning* Educational Review Occasional Publications Number One, pp. 152, 10s. 6d.

Diane Simons (1966). *Georg Kerschensteiner.* London Methuen, pp. 152, 12s. 6d.

Brud J. Biddle and Edwin J Thomas (1966). *Role Theory : Concepts and Research* New York : John Wiley, pp xiv+453, 68s.

Stanley D. Porteus (1966). *Porteus Maze Test : Fifty Years' Application.* California : Pacific Books, pp. 320, 55s.

M. Varma (1966). *An Introduction to Educational and Psychological Research.* London : Asia Publishing House, pp. xi+208, 42s.

W. G. Bennis and E. H. Schein (1966). *Leadership and Motivation.* Mass. : The M.I.T. Press, pp. xxiii+286, 53.

O. A. Oeser (1966). *Teacher, Pupil and Task : Elements of Social Psychology Applied to Education.* Social Science Paperbacks. London : Associated Book Publishers, pp. xiii+196, 25s.

J. A. Keats, H. M. Smith, C. C. Rogers and G. P. Rowe (1966). *Dunwich : A Study of Aboriginal and European Integration.* Australia : University of Queensland Press, pp 47, $0.60.

J. C. Raven (1966). *Psychological Principles Appropriate to Social and Clinical Problems.* London : H. K. Lewis and Co., pp. xx+187, 25s.

G Kalton (1966). *The Public Schools : A Factual Survey.* London : Longmans, pp xxxii+179, 14s.

J. Derrick (1966). *Teaching English to Immigrants* London : Longmans, pp. 256, 12s. 6d.

W. T. Martin and Dan C. Pinck (1966). *Curriculum Improvement and Innovation* : A partnership of students, school teachers and research scholars. U.S.A. : Robert Bentley, Inc., pp. xii+292, $8.95.

M. Tardy (1966). *Le Professeur Et Les Images.* Paris : Presses Universitaires De France, pp. 128, no price.

E. Durkheim (1966). *Education Et Sociologie.* Paris Presses Universitaires De France, pp. 120, no price.

Yvette Hatwell (1966) *Privation Sensorielle Et Intelligence.* Paris : Presses Universitaires De France, pp. xiii+232, no price.

Max Wertheimer (1966). *Productive Thinking.* London · Tavistock Publications, Ltd., pp. xvi+302, 21s.

Myre Sim (1966). *Tutors and their Students.* Edinburgh : E. and S Livingstone, Ltd., pp. 70, 5s.

J. S. Bruner (1966). *Studies in Cognitive Growth.* New York : John Wiley, pp. xvi+343, 60s.

R A Levine (1966). *Dreams and Deeds : Achievement Motivation in Nigeria.* Chicago : The University of Chicago Press, pp. 123, 29s.

O. Lowenstein (1966). *The Senses* Middlesex : Penguin Books, Ltd., pp. 214, 4s.

S. Diamond, (1966). *Information and Error : An Introduction to Statistical Analysis.* New York : Basic Books, Inc., xii+307, 27s. 6d.

R. H. Ojemann (1966). *The School and the Community Treatment Facility in Preventive Psychiatry* : Proceedings of the Fifth Institute on Preventive Psychiatry. Iowa : The University of Iowa, pp. 145, $2.00.

G. A. Miller (1966). *Psychology : The Science of Mental Life.* Middlesex : Penguin Books, Ltd., pp. 413, 7s. 6d.

R. N. Farquhar (1966). *Agricultural Education in Australia.* London : Australian Council for Educational Research, pp. x+322, $6.00.

L. S. Hearnshaw (1966). *The Comparative Psychology of Mental Development.* London : The Athlone Press, University of London, pp. 23, 6s.

P. Groves and L. Stratta (1966). *At the Market, Book* 6. London : Longmans, pp. 47, 6s.

L. Sjodahl (1966). *Elementary Industrial Psychology.* London : Longmans, pp. viii+154, 12s. 6d

P. Groves and L. Stratta (1966). *Sandra Helps the Gang, Book* 5. London · Longmans, pp 47, 6s.

S Cunnison (1966) *Wages and Work Allocation.* London : Tavistock Publications, Ltd., pp. xxvi+291, 42s.

I. J. Gordon (1966). *Studying the Child in School.* New York : John Wiley, pp. xi+145, 38s.

STAN GOOCH and M. L. KELLMER-PRINGLE (1966). *Studies in Child Development: Four Years On.* London: Longmans, pp. xii+157, 17s. 6d

R. V. HECKEL and L. J. PEACOCK (1966) *Textbook of General Psychology.* U.S.A.. The C. V. Mosby Co., Saint Louis, pp. xii+480, no price.

J. BOWLBY (1966). *Maternal Care and Mental Health.* New York: Schocken Books, pp. 357, 22s.

S. COOPERSMITH (1966). *Frontiers of Psychological Research.* San Francisco: W. H. Freeman and Co, pp. viii+322, 36s.

A. N. OPPENHEIM (1966). *Questionnaire Design and Attitude Measurement.* London: Heineman, pp. viii+298, 35s.

EARL A. TAYLOR (1966). *The Fundamental Reading Skill, Second Edit.* U.S.A.: Charles C. Thomas, pp. xviii+157, $.11.50.

P. LEONARD (1966). *Sociology in Social Work.* London: Routledge and Kegan Paul, pp. viii+119, 15s.

P. GROVES and L. STRATTA (1966). *The Club Dance Tempo Book* 4., London: Longmans, pp. 44, 6s.

R. A. BAUER (1966). *Social Indicators.* U.S.A.: The M.I.T. Press, Cambridge, Mass., pp. xxi+357, no price.

C. M. HEINICKE and I. J. WESTHEIMER (1966). *Brief Separations.* London: Longmans, pp. xi+355, 63s.

S. B. SARASON (1966). *Psychology in Community Settings.* New York: John Wiley, pp. x+714, 98s.

R. L. CURLING and A. BLAKELEY (1966). *Echoes of Experience.* London: Longmans, pp. ix+131, 7s. 6d.

N. M. CHANSKY (1966). *Untapped Good: The Rehabilitation of School Dropouts.* U.S.A.: Charles C. Thomas, pp. xviii+267, $9.75.

J. E. MEADE and A. S. PARKES (1966). *Genetics and Environmental Factors in Human Ability.* Edinburgh: Oliver and Boyd, pp. xi+242, 57s. 6d.

G. S. OSBORNE (1966). *Scottish and English Schools: A Comparative Survey of the Past Fifty Years.* London: Longmans, pp. xv+351, 42s.

C. CHERRY (1966). *On Human Communication,* 2nd Edit. U.S.A.: The M.I.T Press, Cambridge, Mass., pp. xiv+337, 80s.

F. SMITH and G. A. MILLER (1966). *The Genesis of Language.* U.S.A.: The M.I.T. Press, Cambridge, Mass., pp. xii+400, 80s.

E. COX (1966). *Changing Aims in Religious Education.* London: Routledge and OKegan Paul, pp. vi+102, 7s. 6d.

S. M. JOURAD and D. C. OVERLADE (1966). *Reconciliation* (Insight 35). U.S.A.: D. Van Nostrand, pp. x+148, 14s.

M. ABRAHAMSON (1966). *Interpersonal Accommodation.* U.S.A.: D. Van Nostrand, pp. vii+150, 14s.

N. J. COLARELLI and S. M. SIEGEL (1966). *Ward H—An Adventure in Innovation* (Insight 33). U.S.A.: D. Van Nostrand, pp. viii+248, 18s.

P. BAKAN (1966). *Attention* (Insight 34). U.S.A.: D. Van Nostrand, pp. vi+225, 14s.

J. W. TIBBLE (1966). *The Study of Education.* London: Routledge and Kegan Paul, pp. x+239, 25s.

J. LYONS and R. J. WALES (1966). *Psycholinguistics Papers.* Edinburgh: University Press, pp. 243, 42s.

H. M. SCHRODER, *et al.* (1966). *Human Information Processing*: Individuals and Groups Functioning in Complex Social Situations. New York: Holt, Rinehart and Winston, pp. viii+224, $7.95.

R. BALBERNIE (1966). *Residential Work With Children.* Oxford: Pergamon Press, pp. x+244, 45s.

J. A. CLARK, and C. H. DOBINSON (1966). *Earth and Universe.* London: Longmans, pp. 96, 5s.

D. W. WARWICK (1966). *The Advent of Our Faith—The Living Bible,* Book 3. London: Longmans, pp. 103, 7s. 6d.

F. W. MITCHELL (1967). *Sir Fred Clarke—Master-Teacher* 1880-1952. London: Longmans, pp. xvii+236, 45s.

M. L. Kellmer-Pringle (1967). *Studies in Child Development—11,000 Seven-year-Olds.* London. Longmans, pp. xx+246, 21s.

M. L. Kellmer-Pringle (1967). *Studies in Child Development—Adoption Facts and Fallacies.* London: Longmans, pp. x+251, 21s.

J. M Morris (1966). *Standards and Progress in Reading.* Slough: The National Foundation for Educational Research in England and Wales, pp. xxxiv+493, 50s.

E. Anstey (1966). *Psychological Tests.* London: Nelson, pp. xix+300, 63s.

J D. Halloran (1967). *Attitude Formation and Change.* Leicester University Press, pp. 167, 21s

D. H. Harding (1967). *The New Pattern of Language Teaching.* London: Longmans, pp. viii+212, 12s. 6d.

M. G. Frankel, F. Wm. Happ and M. P. Smith (1966). *Functional Teaching of the Mentally Retarded.* U.S A.: Charles C. Thomas, pp. xvii+241, $9.75.

R. D. Laing (1967). *The Politics of Experience and the Bird of Paradise.* Penguin Books, Ltd, pp. 155, 4s. 6d.

J Chateau (1966). *Les Grands Pedagogues.* Paris: Presses Universitaires De France, pp. xii+362, no price.

J. Meltzoff and R. L. Blumenthal (1966). *The Day Treatment Centre.* U.S.A.: Charles C. Thomas, pp. xvi+234, $8.75.

M. Craft, J. Raynor and L. Cohen (1967). *Linking Home and School.* London: Longmans, pp. x+246, 27s. 6d.

S. D. Porteus (1965). *Porteus Maze Test*: Fifty Years' Application. California: Pacific Books, pp. 320, $7.50.

N. D. Rothwell (1966). *The Psychiatric Halfway House*: *A Case Study.* U.S.A.: Charles C. Thomas, pp. xiv+265, $8.75.

P. H. Berkowitz and E. P. Rothman (1966). *Public Education for Disturbed Children in New York City*: Application and Theory. U.S.A.: Charles C. Thomas, pp. x+376, $12.50.

Gordon Rose (1967). *Schools for Young Offenders.* London: Tavistock Publications, pp. ix+244, 52s.

G Mandler, P. Mussen (1967). *New Directions in Psychology—*III. New York: Holt, Rinehart and Winston, pp. viii+289, $3.95.

M. Sherif (1967). *Group Conflict and Co-operation.* London: Routledge and Kegan Paul, pp. xiii+192, 21s.

C Frankenstein (1967). *The Roots of the Ego.* Edinburgh: E. and S. Livingstone, pp. xi+286, 62s.

D. A. Schon (1967). *Invention and the Evolution of Ideas.* London: Tavistock Publications, pp. xiii+208, 21s.

J. Hillman (1967). *Insearch*: *Psychology and Religion.* London: Hodder and Stoughton, pp. 116, 21s.

G. W. H. Gosling (1967). *Marking English Compositions.* London: Australian Council for Educational Research, pp. x+94, $2.50.

A. Workbook (1967). *Elementary Statistics*: *A Workbook.* Oxford: Pergamon Press, pp. ix+100, 15s.

S. Hegeler (1966). *Peter and Caroline.* London: Tavistock Publications, pp. v+29, 6s.

R. S. Peters (1967). *The Concept of Education.* London: Routledge and Kegan Paul, Ltd., pp. viii+223, 30s.

P. L. Ewalt (1967). *Mental Health Volunteers.* Illinois: Charles C. Thomas, pp. xxiii+160.

J. E. Cass (1967). *Literature and the Young Child.* London: Longmans, Green and Co., Ltd., pp. xv+109, 9s. 6d.

K. Foppa (1966). *Lernen Gedachtnis Verhalten*: Koln: Kiepenheuer and Witsch, pp. 466, DM24,80.

D H. Hargreaves (1967). *Social Relations in a Secondary School.* London: Routledge and Kegan Paul, Ltd., pp. xii+226, 32s.

M. Sturt (1967). *The Education of the People.* London. Routledge and Kegan Paul, Ltd., pp.xi+432, 60s.

R. Kaelbling and R. M. Patterson (1966). *Eclectic Psychiatry* Illinois : Charles C. Thomas, pp. xxv+891, $17 50.

D. Unwin and J. Leedham (1967). *Aspects of Educational Technology.* London . Methuen and Co., Ltd., pp vii+545, 5 gns.

S. Cunnison (1966). *Wages and Work Allocation.* London : Tavistock Publications, pp. xxvi+291, 42s.

E. Venables (1967). *The Young Worker at College.* Bristol : Western Printing Services, Ltd., pp. 240, 36s.

M. Raychaudhuri (1966). *Studies in Artistic Creativity.* Calcutta : Rabindra Bharati, pp. xii+256, Rs. 15.00

S. J. Armore (1967). *Introduction to Statistical Analysis and Inference . . for Psychology and Education.* London : John Wiley and Sons, pp. xx+546, 72s.

H. Sinclair-de Zwart (1967). *Acquisition Du Langage et Developpement De La Pensee.* Paris : Dunod, pp. vi+168, 23F.

G. R. Medinnus (1966). *Reasdings in the Psychology of Parent-Child Relations.* London : John Wiley and Sons, Ltd., pp. xiii+371, 35s.

B. Way (1966). *Development through Drama.* London Longmans, Green and Co., Ltd., pp. viii+308, 15s. 6d.

T. F. Graham (1967) *Medieval Minds.* London : George Allen and Unwin, Ltd., pp. 112, 30s.

V. Ruke-Dravina (1967). *Mehrsprachigkeit Im Vorschulalter.* Sweden : C. W. K. Gleerup, pp. 104, 15 couronnes sued.

W. R. Jones (1966). *Bilingualism in Welsh Education.* Cardiff : University of Wales Press, pp. xiv+202, 18s

University of Bristol Institute of Education (1967). *The Assessment and Education of Slow-Learning Children.* London : University of London Press, Ltd., pp. 72, 7s. 6d.

J. B. Mays (1967). *The School in its Social Setting.* London : Longmans, Green and Co., Ltd., pp. x+94, 10s. 6d.

Hertfordshire Social Studies Committee (1967). *Teaching Liberal Studies.* London . Longman's Green and Co , Ltd., pp. vii+128, 15s.

F. G. Lennhoff (1967). *Being Sent Away.* Harper Hill : A Shotton Hall Publication, pp. 72, 6s.

D. A. Pidgeon (1967). *Achievement in Mathematics.* Hove . King, Thorne and Stace, Ltd., pp. vi+317

AN INTRODUCTION TO EDUCATIONAL PSYCHOLOGY

E. Stones

" Here, at last, is the book we have been waiting for, an introduction to educational psychology which links theory to practice, and which draws on the most advanced thinking and recent scientific developments." Brian Simon, *Forum* **36s**

University Paperback **18s**

CONTRARY IMAGINATIONS

A PSYCHOLOGICAL STUDY OF THE ENGLISH SCHOOLBOY

Liam Hudson

" There are a few, a very few, books on educational psychology which grip from the start, which speak without protective coatings of jargon, which genuinely open fresh doors. Dr. Hudson has without doubt written one. *Contrary Imaginations* is full of fresh insights, freshly expressed." Peter Preston, *The Guardian* **25s**

THE PSYCHOLOGY OF CHILDHOOD AND ADOLESCENCE

C. I. Sandström

" This is a remarkable little book, containing much valuable information . . . One can confidently recommend for reading to teachers, educationalists and undergraduates taking their first steps in the field of developmental psychology, and also to those who are seeking to revise quickly and easily their knowledge of this subject." *Mental Welfare* **36s**

READINGS ON THE EXCEPTIONAL CHILD

Edited by **E. P. Trapp and P. Himelstein**

These readings cover the field of the mentally deficient, the exceptionally gifted, the physically handicapped and the emotionally disturbed child **55s**

PRINCIPLES OF PSYCHOLOGICAL MEASUREMENT

G. C. Helmstadter

" Certainly a well-written and logical text like the one under review is greatly to be preferred to the volumes on measurement which devote a high percentage of their content to detailed descriptions of sample tests and make little or no attempt to deal with basic ideas and principles." *Personnel and Guidance Journal* **36s**

METHUEN

THE DEVELOPMENT OF SEX DIFFERENCES
Edited by **Eleanor E. Maccoby**

This book deals with differences between the sexes in behaviour other than specifically sexual behaviour. The focus is on how sex differences in temperament and ability develop in young children, and on possible explanations for these differences.

Stanford Studies In Psychology **60s net**

EXPERIENTIAL FOUNDATIONS OF RORSCHACH'S TEST
Ernest G. Schachtel

This book makes a highly important contribution to the theory and practice of the Rorschach method. In combining the experiential-phenomenological and psychoanalytic views of the subject's responses to the Rorschach ink blots, Dr Schachtel reveals a new and significant dimension in the processes of the individual as he takes the test.

Studies in Existentialism and Phenomenology **55s net**

SCHOOLS FOR YOUNG OFFENDERS
Gordon Rose

' Dr Rose provides here quite the finest account available of the development, the structure and the functioning of approved schools in England and Wales.'

The Times Educational Supplement

International Library of Criminology No. 17 **52s net**

ABSENT
School Refusal as an Expression of Disturbed Family Relationships

Max B. Clyne

' This book is addressed to general practitioners, teachers and others who work with children. It gives a very full account of the problems of school refusal and it is rewarding to examine the case records of one practitioner whose experience of this often complicated disorder is extensive.' Edmund A. Harvey-Smith, *World Medicine*

' This is a useful and readable medical-sociological study.'

The Times Educational Supplement

38s net

A STUDY OF CHILDREN'S THINKING
Margaret Donaldson

'. . . anyone interested in children's thinking should make a careful study of the contents of this book, which provide a fascinating insight into the difficulties which children encounter in developing from simple concrete to abstract logical thinking.'

The Times Educational Supplement

35s net

Now available in paperback

TEACHER, PUPIL, AND TASK
Elements of Social Psychology applied to Education

Edited by **Oscar A. Oeser**

' . . . Professor Oeser has succeeded in producing the most useful and comprehensive treatment of the topic that has been published to date.'

The Times Educational Supplement

hardback 25s net
paperback 16s net

TAVISTOCK PUBLICATIONS

Spring Grove

The Education of Immigrant Children

TREVOR BURGIN *and* PATRICIA EDSON

'*Spring Grove* carries a deeper significance, and makes a greater contribution to the understanding of the growing multiracial character of Britain than it claims to do as the documentation of an educational experiment . . . . It is an exhilarating essay on integration as we seek to define the term as an interpretation of our social attitudes towards our new citizens.' **The Teacher** 21*s net* *Institute of Race Relations*

The Lore and Language of Schoolchildren

IONA *and* PETER OPIE

'It is the sort of book one can brood over with continually renewed nostalgia, amusement, and wonder.' *Richard Hoggart in* **The Guardian**

'As well as being a valuable social study, it is one of the most exhilarating anthologies of our day.' **The Times Literary Supplement** 11 *text figures* 12*s* 6*d net* *Oxford Paperbacks*

Special Education in England and Wales

STEPHEN JACKSON

The study of Special Education now forms part of the training of most teachers, many of whom are likely to find themselves teaching at least a few children who are to some degree handicapped, physically, emotionally, or educationally. This book covers a wide range of problems and should be of interest to parents as well as teachers. 13*s* 6*d net*

The College Dropout and the Utilization of Talent

LAWRENCE A. PERVIN, LOUIS E. REIK
and WILLARD DALRYMPLE

This book, written by a group of social scientists, educators and psychiatrists, examines the complicated personal and institutional factors that influence a student's decision to break off his studies, and investigates methods of enabling him to resume his education and utilize his talents. 3 *text figures* 52*s net* *Princeton University Press*

OXFORD UNIVERSITY PRESS

AUSTRALIAN JOURNAL OF PSYCHOLOGY

Editor : I. K. WATERHOUSE, Macquarie University, N.S.W.

Vol. 18 *No.* 3 *December*, 1966

S. B. Hammond. An Interactionist Approach to Personality.
G. E. O'Brien. Behaviourism and Explanation.
Anthony R. Stone, Stanley D. Imber and Jerome D. Frank. The Role of Non-Specific Factors in Short-Term Psychotherapy.
C. C. Anderson and A. J. Cropley. Some Correlates of Originality.
J. A. Radcliffe. WAIS Factorial Structure and Factor Scores for Ages 18 to 54.
Wendy L. Walker. A Preliminary Investigation of Personality Change in Depressed Patients Following ECT.
P. N. Russell and R. A. M. Gregson. A Comparison of Intermodal and Intramodal Methods in the Multidimensional Scaling of Three-Component Taste Mixtures.
J. J. Furedy. Reaction Time as an Index of Masking and the Effect of Check Trials on Thoughtful Subjects.
R. Over. Context and Movement as Factors Influencing Haptic Illusions.
Gordon Stanley and J. J. Furedy. Size Constancy and Emmert's Law of Apparent Sizes.
G. E. Kearney. Hue Preferences as a Function of Ambient Temperatures.
A. Ivinskis. A Note on the Open-Field Test of Emotionality.
Aubrey J. Yates. A Note on Progressive Matrices (1962)
R. J. Harrison. Predictive Validity of Some Adjustment Tests : a Research Note.
I. F. Paton. A Note on a Simple Method of Making Lip-Synchronized Sound Film.

The Journal is published three times a year. Subscription rates 36s. (sterling), and $5·50 in U.S.A. Members of B.P.S. and A.P.A. should enquire about special subscription rates from the Secretary of their Society. Non-members should send enquiries to The Executive Officer, Australian Psychological Society, Department of Psychology, University of Sydney, Sydney, N.S.W., Australia.

A Teacher's Guide to Reading Piaget

MOLLY BREARLEY AND ELIZABETH HITCHFIELD

'Deserves to succeed in its object of giving teachers an *apéritif* prior to reading Piaget in the raw. The authors have confined themselves to looking at eight specific aspects of Piaget's work : number, measurements, knots, perspective, co-ordinates, floating and sinking, moral judgment, and the behaviour of babies, and they wisely adopt the policy of describing an experiment prior to discussing its theoretical implications. What the writers have to say will be particularly useful to teachers in nursery and primary schools.'—*Teachers World.* 18*s.*

Adolescence

Its Social Psychology

C. M. FLEMING

'This is an important and valuable book. It sums up recent research on the physiology and psychology of adolescence, as well as its educational and social aspects. Doctors, teachers, sociologists and psychologists will all find much to stimulate thought. Presented as it is, without prejudice or bias, it is to be highly recommended'—*British Medical Journal* 25*s.* *Paperback* 12*s* 6*d.*

ROUTLEDGE & KEGAN PAUL

The BRITISH JOURNAL OF PSYCHOLOGY

EDITED BY ARTHUR SUMMERFIELD

CONTENTS OF VOLUME 57, PART 3-4, 1967

LEVELT, W. J. M. The alternation process in binocular rivalry.
DIXON, N. F. EEG correlates of threshold regulation as a function of stimulus wavelength. a comparison between normal subjects and psychiatric patients.
TAYLOR, ANN and FORREST, DEREK W. Two types of set and the generalization of perceptual defence.
WICKELGREN, WAYNE A. Numerical relations similarity, and short-term recognition memory for pairs of digits.
LEONARD, J. A The development of a keyboard training situation in miniature.
TAYLOR, FITZ J. The effect of learning instruction on motor and verbal responses in recall.
WETHERICK, N. E. The responses of normal adult subjects to the matrices test.
CAINE, T. M. Perceptual sensitization, response suppression, and psychopathology.
HOTOPF, W. H. N. The size-constancy theory of visual illusions.
HAMILTON, VERNON. Size constancy and intelligence ; a re-examination.
STANNARD, R., SINGER, G. and OVER, R The effect of informational feedback on size judgements of schizophrenic patients.
OVER, RAY. An experimentally induced conflict between vision and proploception.
BATTIG, WILLIAM F. Effects of sequential ordering of added pairs and of correction V Non-correction procedures on paired associated performance.
CHEYNE, WILLIAM M. Vanishing cues in paired-associate learning
CLARKE, ANN M. and COOPER, G. M. Transfer in category learning of young children : its relation to task complexity and overlearning.
KEEHN, J. D. Avoidance responses as discriminated operants.
DAVIES, D. R and HOCKEY, G. R. J. The effects of noise and doubling the signal frequency on individual differences in visual vigilance performance.
TUNE, G. S Age differences in errors of commission.
McLAUGHLIN, R J. and EYSENCK, H. J Visual masking as a function of personality.
ALKER, HENRY A. A psycholinguistic investigation of a rule theory of ' sense '
JONES, SHEILA. Decoding a deceptive instruction.
WASON, P. C. and KOSVINER, ADELE. Perceptual distortion induced by reasoning.
EYSENCK, H. J. A reply to McPherson's comments on the Aberdeen conditioning studies.
STOTT, D. H. Commentary on ' the genetic determination of differences in intelligence: a study of monozygotic twins reared together and apart ' by Cyril Burt : congenital influences on the development of twin.
LEWIS, D. G. Commentary on ' the genetic determination of differences in intelligence . a study of monozygotic twins reared together and apart ' by Cyril Burt.
HOWE, M. J. A. A note on order of recall in short-term memory.
PUBLICATIONS REVIEWED.

Subscriptions 90s. net per volume of 4 parts
Orders may be placed with a bookseller or sent direct to the publishers
CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200 Euston Road, London, N.W.1

THE BRITISH JOURNAL OF MEDICAL PSYCHOLOGY

Edited by Dr. THOMAS FREEMAN

Volume 40, *Part* 2. CONTENTS *June*, 1967

R. D. SCOTT and P. L. ASHWORTH. ' Closure ' and the First Schizophrenic Break-Down.
SAMUEL D. LIPTON. Freud's Position on Problem Solving in Dreams.
G. A. FOULDS and P. R. MAYO. Neurotic Symptoms, Intropunitiveness and Psychiatric Referral.
PAUL KLINE. Obsessional Traits and Emotional Instability in A Normal Population.
BURTON A. GLICK. A Note on Freud's Empty Neurone Theory.
BRUCE DENNER and ROGER BIBACE. A Developmental Analysis of the Amnesic Syndrome.
J. W. WARBURTON. Memory Disturbance and The Parkinsonian Syndrome.
D. C. KENDRICK. A Cross-Validation Study of the Use of the SLT and DCT in Screening for Diffuse Brain Pathology in Elderly Subjects.
T. M. CAINE and D. J. SMAIL. Attitudes of Psychiatrists to Staff Rôles and Treatment Methods.
HARVEY R. GREENBERG and NOEL LUSTIG. Remission of Malingered Amnesia Following a Natural Disaster.
Reviews.

Subscription £4 10*s. per volume of* 4 *parts. Single parts* 30*s. plus postage*

Orders may be placed with any bookseller, or direct to :

CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200, Euston Road, London, N.W.1

Studies in Child Development

This series of books, published by Longmans, in association with the National Bureau for Co-operation in Child Care, has a twofold object : to make the results of research available to a wide public as rapidly as possible ; and to disseminate, with a minimum of technical jargon, research findings concerned with the development and needs of children.

FAMILY ADVICE SERVICES:

An exploratory study of a sample of such services organised by Children's Departments in England

Aryeh Leissner *12s. 6d.*

FOUR YEARS ON

S. Gooch & M. L. Kellmer Pringle

17s. 6d.

11,000 SEVEN-YEAR-OLDS

M. L. Kellmer Pringle, N. R. Butler and R. Davie *21s.*

ADOPTION–FACTS & FALLACIES

M. L. Kellmer Pringle

with the assistance of **E. Crellin, R. Dinnage** *and* **M. Dewdney** *21s.*

Longmans

NEW BOOKS

Statistical Methods in Education

D. G. LEWIS, M.SC., B.ED., PH.D.

Lecturer in Methods of Educational Research, University of Manchester

Books on statistical methods usually make great demands on the mathematical competence of their readers. In this book, however, no mathematical knowledge is presumed beyond a knowledge of the basic rules of arithmetic.

The author deals with the fundamental concepts of descriptive statistics—means, standard deviations, correlation, regression, test norms including I.Q.s, etc.,—and the process of generalizing from the sample to the group from which the sample is taken.

He develops sampling theory to include the t, F and chi-squared tests of significance and the analysis of variance (simple and complex cases), and discusses the reliability and validity of tests. Illustrations from the field of education introduce each of the methods considered.

The book is intended for students of educational psychology, and especially for those doing experimental work. It should be of particular interest to those reading for the M.Ed. and M.A. (Educ.) degrees and to students working for the B.Ed. degree. 35/- net.

Two books first published by University of Toronto Press

The Child and the Institution

A study of deprivation and recovery

BETTY M. FLINT

This is the report of a study of small children who had spent their first few years in an institution. The enquiry was made jointly by the Catholic Children's Aid Society of Toronto and the Institute of Child Study of the University. The Catholic Children's Aid Society wished to find out how to cure or relieve the effects of deprivation already suffered by children living in an institution and how to prevent this kind of deprivation in future. The Institute of Child Study was contemplating research into the way children's early development is affected by the lack of normal care and affection. 30/- net.

Human Security

W. E. BLATZ

Professor Blatz's Theory of Security was very familiar to a generation of students at the University of Toronto, but it was only towards the end of his life that he wrote it down in book form.

The Theory of Security is based on a unifying principle which Professor Blatz called 'security' and defined as 'the state of mind which accompanies the willingness to accept the consequences of one's acts—without equivocation of any sort.' This book was written to describe the development of security patterns in human beings. 25/- net.

Order these books through your bookseller

UNIVERSITY OF LONDON PRESS LTD

St. Paul's House, Warwick Lane, London, E.C.4

OCCUPATIONAL PSYCHOLOGY

Editor : ALEC RODGER

JANUARY 1967 VOLUME 41, No. 1

CONTENTS

'Technical' Change and the Misuse of Professional Manpower · Some Studies and Observations. By DENIS PYM.

Why Ergonomics? By K. F H. MURRELL.

Comments on Murrell's 'Why Ergonomics?' By W. T. SINGLETON, B. SHACKEL, K. F. JACKSON and E. N. CORLETT.

Extending the Occupational Environment : The Measurement of Organisations. By KERR INKSON, ROY PAYNE and D. PUGH.

The Evaluation of Visual Displays : Techniques and Trends By J. M. ROLFE.

Isolation and Description of Occupational Stereotypes. By JOAN GUILFORD.

Book Reviews.

Annual Subscription 40 *shillings, overseas* 50 *shillings*

National Institute of Industrial Psychology

14, Welbeck Street, London, W.1.

The British Journal of
Social and Clinical Psychology

Edited by MICHAEL ARGYLE and R. R. HETHERINGTON

Contents of Vol. 6, Part 1, February, 1967

FISHER, R M. Acquiescent response set, the Jesness Inventory and implications for the use of 'foreign' psychological tests.

LEACH, P. J. A Critical study of the literature concerning rigidity.

WRIGHT, D. Religious belief and co-education in a sample of sixth-form boys and girls.

McINTYRE, D and MORRISON, A. The educational opinions of teachers in training.

CAINE, T. M Response consistency and testing levels

WALTON, H J. and HOPE, K. The effect of age and personality on doctors' clinical preferences.

FOULDS, G. A. Some differences between neurotics and character disorders.

VENABLES, P. H. The relation of two flash and two click thresholds to withdrawal in paranoid and non-paranoid schizophrenics.

MAYO, P. R. Some psychological changes associated with improvement in depression.

ARMSTRONG, Jnr, H. E., JOHNSON, M. H., RIES, H. A. and HOLMES, D. S. Extraversion—introversion and process-reactive schizophrenia.

KIPPER, D A. Diagnoses of psychotic patients with the Kahn Test of symbol arrangement

McDONALD, K. G. and CROOKES, T. G. The WAIS picture arrangement test in British psychiatric patients.

BOOK REVIEWS.

The subscription price per volume of four parts is 90s. net. Orders may be sent to your bookseller or direct to the publishers

CAMBRIDGE UNIVERSITY PRESS

Bentley House, 200, Euston Road, London, N.W.1

OTHER PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology.
The British Journal of Psychology (Monograph Supplements).
The British Journal of Medical Psychology.
The British Journal of Mathematical and Statistical Psychology.
The British Journal of Social and Clinical Psychology.

The British Journal of Psychology and *The British Journal of Medical Psychology* appear in quarterly parts in volumes of about 350 pages, subscription price of each, payable in advance, 60s. per volume (post free). Subscriptions may be sent to any Bookseller or to the Cambridge University Press, Bentley House, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $10.00. Enquiries should be addressed to the Cambridge University Press, American Branch, 32, East 57th Street, New York 22.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November, in volumes of about 184 pages. The subscription rate payable in advance is 30s. 6d. per volume (post free). The subscription price in the U.S.A. is $4.50 (post free). Subscriptions should be sent to the Society's Agents, Taylor and Francis, Ltd., Red Lion Court, Fleet Street, E.C.4, or to any Bookseller.

The British Journal of Social and Clinical Psychology is issued three times a year in February, June and November, in volumes of approximately 240 pages. The subscription price of each, payable in advance, is 50s. (post free). Subscriptions may be sent to any Bookseller, or to the Cambridge University Press, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $8 50. Enquiries should be sent to the Cambridge University Press, American Branch, 32, East 57th Street, New York 22. Separate parts will be sold at 20s. : $3.50 in the United States.

A series of *Monograph Supplements* is issued in connection with the *British Journal of Psychology*, but is not included in the subscription for that journal. Orders may be sent to any Bookseller or to the Cambridge University Press, 200, Euston Road, London, N.W.1, from whom details of these publications may be obtained.

MANUSCRIPTS

For the *British Journal of Psychology* should be sent to Professor ARTHUR SUMMERFIELD, Department of Psychology, Birkbeck College, Malet Street, London, W.C.1.

For the *British Journal of Medical Psychology*, should be sent to Dr. THOMAS FREEMAN, Lansdowne Clinic, 4, Royal Crescent, Kelvingrove, Glasgow, C.3.

For the *British Journal of Educational Psychology*, should be sent to Mr. L. B. BIRCH, Institute of Education, The University, Sheffield, 10.

For the *British Journal of Mathematical and Statistical Psychology*, should be sent to Dr. R. J. AUDLEY, Department of Psychology, University College, Gower Street, London, W.C.1.

For the *British Journal of Social and Clinical Psychology*, should be sent to Mr. MICHAEL ARGYLE (Social Psychology Editor), Institute of Experimental Psychology, 1, South Parks Road, Oxford;

or to:

Dr. RALPH HETHERINGTON (Department of Studies in Psychological Medicine), 6, Abercromby Square, Liverpool, 7.

For the *Monograph Supplements*, to Professor IAN M. L. HUNTER, Department of Psychology, University of Keele, Keele, Staffs., whose decision regarding the publication of all *Monograph Supplements* will be final. No MS. can be considered for publication in any year unless it is in the hands of the Editor by 1st September of the preceding year.

OFF-PRINTS

Contributors receive 50 free copies of their article and may buy more in units of 50.

Members of the British Psychological Society receive all the Society's Journals at reduced rates.

Information on membership and further activities may be obtained from the Secretary, Tavistock House South, Tavistock Square, London, W.C.1.

PUBLISHED BY THE ASSOCIATION OF TEACHERS IN COLLEGES AND DEPARTMENTS OF EDUCATION

Education for Teaching is published three times a year at mid-term in November, February and May, and distributed free to members of the Association. To non-members it costs 5s. per issue, post free (subscription 14s. per annum).

Members of the A.T.C.D.E. receive the *British Journal of Educational Psychology* at a reduced rate.

Information on membership and further activities may be obtained from the Honorary Secretary, 151, Gower Street, London, W.C.1.

8 NOV 1967

CONTENTS

A. Davies : The English Proficiency of Overseas Students 165
E. A. Lunzer and I. Hulme : Discrimination Learning and Discrimination Learning Sets in Subnormal Children 175
J. D. Nisbet and N. J. Entwistle : Intelligence and Family Size, 1949-1965 188
R. K. Turner, A. Mathews and S. Rachman : The Stability of the WISC in a Psychiatric Group . 194
K. Lovell and J. B. Shields : Some Aspects of a Study of the Gifted Child 201
R. B. Cattell : The Theory of Fluid and Crystallized General Intelligence Checked at the 5–6-year-old Level 209
H. A. S. Zahran : The Self-Concept in the Psychological Guidance of Adolescents 225
C. E. Buxton : Evaluations of Three Factors that Might Influence Responses to Tests of Academic Achievement Motivation . . 241

Research Notes :

S. Wiseman : The Effect of Restriction of Range Upon Correlation Coefficients 248
W. Yule : Predicting Reading Ages on Neale's Analysis of Reading Ability 252
F. H. Farley : Birth Order, Achievement-Motivation and Academic Attainment 256
H. G. Armstrong : Wastage of Ability Amongst the Intellectually Gifted . 257
G. H. McLaughlin : Comments on " An Investigation of the Craig Reader Course," by K. D. Duncan and R. Foster 260

Theses Abstracts :

H. H. Richards : Psychological Factors Associated with the Sociometric Status of Children Attending a Comprehensive School in Breconshire 261

Book Reviews 263
Publications Received 276

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

NOVEMBER 1967

VOL. XXXVII, PART 3

Price One Pound

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by
L. B. BIRCH

Associate Editors:
P. E. VERNON and DORIS M. LEE

Advised by

SIR CYRIL BURT
R. A. C. OLIVER
E. A. PEEL
SIR FRED J. SCHONELL
S. WISEMAN

Vol. XXXVII, Part 3

The British Journal of Educational Psychology is issued by the British Psychological Society and the Association of Teachers in Colleges and Departments of Education, once a term, in February, June and November. The subscription is £2 a year post free, payable in advance, or £1 each number.

Subscriptions and Orders should be sent to "The Manager, B.J.E.P. Department, Methuen and Co., Ltd., 11, New Fetter Lane, London, E.C.4. Members of the British Psychological Society and members of the A.T.C.D.E. receive the *Journal* on special terms through the Societies.

Papers for publication should be sent to Professor J. D. Nisbet, King's College, Aberdeen. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for the *British Journal of Educational Psychology* should be sent to Professor J. D. Nisbet, at the above address. *Advertisements* and *Subscriptions* for this *Journal* should be sent to Manager, B.J.E.P. Department, Sweet and Maxwell, Spon Booksellers, Ltd., 11, New Fetter Lane, London, E.C.4.

Published by
METHUEN AND CO., LTD., 11, NEW FETTER LANE, LONDON, E.C.4,
for the British Journal of Educational Psychology, Ltd.

Registered Offices: 42, Hill Street, Birmingham, 5. This address should *not* be used for letters or books intended for the Editor, Authors, or the Publishers.

On behalf of
The British Psychological Society and the Association of Teachers in Colleges and Departments of Education.

All rights reserved.

PRINTED BY THE BIRMINGHAM PRINTERS, LTD., 42, HILL STREET, BIRMINGHAM, 5.

UNIVERSITY

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by

L. B. BIRCH

Associate Editors:

P. E. VERNON and DORIS M. LEE

Advised by

SIR CYRIL BURT
R. A. C. OLIVER
E. A. PEEL
SIR FRED J. SCHONELL
S. WISEMAN

Vol. XXXVII
1967

Published by
METHUEN AND CO., LTD., 36, Essex Street, Strand, London, W.C.2.,
for the British Journal of Educational Psychology, Ltd.
Registered Offices: 42, Hill Street, Birmingham, 5. This address should *not* be used for letters or books intended for the Editor, Authors, or the Publishers.

On behalf of
The British Psychological Society and the Association of Teachers in Colleges and Departments of Education.

All rights reserved.

PRINTED BY THE BIRMINGHAM PRINTERS, LTD., 42, HILL STREET, BIRMINGHAM, 5.

THE JOURNAL

During my own spell as Editor of this *Journal* (1956-61), I was not infrequently abroad, and was fortunate then in being able to rely on Mr. L. B. Birch as Assistant Editor, to carry out the work of getting it to press. Mr. Birch had also long been associated with Professor Valentine, the founder of the *Journal* and Editor for twenty-five years ; hence the Directors had no hesitation in entrusting it to him when I retired. Now, after another five-and-a-half years, he has been offered and accepted the Chairmanship of Postgraduate Studies in the Faculty of Education of McGill University, and is leaving for Montreal forthwith. I am sure that authors and readers who know him will join the Directors in wishing him and his family every success in their new environment, much as we regret his loss to English education. We thank him very warmly for his services to the *Journal*, given so cheerfully and so efficiently over many years. It is hardly necessary to point out that the *Journal* has flourished under him. As the number of subscribers has grown, so has the size of the *Journal*, and at least as high standards of scholarship, experimentation and style of presentation, have been maintained. The 1956 volume contained 242 pages, in 1961 there were 318, and 1967 will approach or exceed 400 pages. Though the price has had to be increased because of tremendous rises in printing and other costs, the *Journal* can still claim to be the best bargain among professional periodicals in any English-speaking country. And the Editor, together with the Treasurer, is responsible for this.

The Directors feel that they are fortunate in being able to obtain the agreement of Professor J. D. Nisbet, Professor of Education in the University of Aberdeen, to take over the Editorship. His colleague, Mr. Ralph Dutch, will act as Assistant Editor. Running a journal of this type has become a more and more complex affair ; moreover, Educational Psychology itself has grown considerably in scope and is being studied by more and more people. Thus, over the next few years there may well be alterations in the make-up and other features of the *Journal* and in Editorial policy. We are confident that Professor Nisbet, while respecting the traditions that Professor Valentine laid down in the past, will institute those changes that are needed for the future.

P. E. VERNON.

ADMINISTRATION OF GROUP INTELLIGENCE TESTS TO EAST AFRICAN PUPILS

By PHILIP E. VERNON

(University of London, Institute of Education)

SUMMARY. Intelligence tests could be even more valuable for educational predictions in the developing, than in Western, countries, because of the large variations in efficiency and linguistic background in different primary schools. But there may be serious problems in getting them across to unsophisticated pupils, especially with printed instructions only.

An experiment was carried out with fourteen top primary classes in Tanzanian schools, where a short non-verbal battery was given with printed, or minimal oral, instructions in English or Swahili. A second parallel form was taken after this practice or after various types of coaching. The mean improvement varied considerably in different classes and different subtests, and was affected by certain motivational or attitude factors. Oral instructions gave better mean scores than printed, and Swahili better than English The overall rise with practice averaged 3 to 4 points in deviation I.Q. units, with coaching about 7 points. These figures are not unusually large, possibly because the pupils were already sophisticated to objective attainment tests. Thus they suggest that intelligence tests could be introduced with the same kind of familiarization procedures common in England, though the actual items should be constructed locally. Comparison of item difficulties with those found in England showed that abstract non-verbal items may be perceived differently in different cultural groups.

I.—INTRODUCTION.

ENTRY to secondary schools in the East African countries of Kenya, Uganda and Tanzania is as competitive as in the days of our own Special Place examination. Approximately 25 per cent. of the population complete the seven standards of primary schooling at ages ranging from 11 years to over 20 (the majority at 13 to 15), and there are secondary places available for 10–12 per cent. of these. Moreover, the pressure is mounting as greater numbers enter and complete primary schooling, and the quality of primary teaching improves. Admission is based at present mainly on objective or semi-objective tests in English, Arithmetic and (in Tanzania) General Knowledge and Swahili. Not only do these stimulate cramming by the schools, but also the results vary much more widely from one school to another than in the U.K., depending on the region of the country—its linguistic background and its cultural and economic advancement, or on the quality of the staff and particularly on their English teaching (cf. Irvine, 1965). Hence, there is an even stronger case for the introduction of tests of the intelligence type than there was in U.K. in the 1920's, in the hope of reducing school differences, dependence on cramming and variations in cultural background. Although we admit that these can never be culture-free, even when based on pictures or abstract materials (cf. Vernon, 1966b), they should give better predictions of aptitude for advanced education than the present system (Silvey, 1964).

However intelligence testing raises much more difficult problems of two types than in U.K., difficulties of content and of administration. All these countries are linguistically heterogeneous ; English is introduced as the medium of instruction quite early in most Kenyan and Ugandan schools, not till fifth or sixth standard in most Tanzanian schools where Swahili is regarded as the national *lingua franca*. Thus, there is no really suitable common medium on which to base test items. Group tests using non-verbal materials have been constructed, e.g., by Oliver (1932) and Lule, though these too are likely to vary in comprehensibility in different tribal areas, and are anyhow known to give poorer predictions of academic educability than verbal materials. Careful experimentation is needed to discover the simplest and most adaptable item-types, whether English, Swahili, pictorial or abstract, for example, following the lead given by Scott in the Sudan (1950).

Administrative difficulties arise from the total unfamiliarity of such tests to East Africans. The analogy, the odd-man-out, the multiple-choice item, even working on one's own at printed materials with a time limit, may all be unnatural to the modes of thought and attitudes of non-western cultures. Biesheuvel (1949) and Schwarz (1961) have indeed shown that group tests can be got across to, and yield reliable and valid results with, illiterate African adults; the former by using sound-films of an African tester who explains every step of the procedure, the latter by intensively teaching the testees how to do each test before they begin. But such techniques would be quite inapplicable when thousands of children have to be tested simultaneously by untrained invigilators (mostly primary heads or secondary teachers) all round the country. Thus it would be essential to use printed instructions only, though steps could be taken to promote some previous familiarization in the schools.

It seemed desirable, therefore, to explore the extent of practice and coaching effects before attempting to construct usable tests One previous investigation of this problem has been carried out in South Africa by Lloyd and Pidgeon (1961), with 10½–12½-year African, Indian and European children. Table 1 shows the initial mean I.Qs. of representative samples of pupils on the N.F.E.R. Nonverbal Tests 1 and 2, the gains resulting from practice (i.e., taking a second test a fortnight later), and the additional gain when two half-hour periods of coaching in English or the veranacular were inserted before the second test. Note the large practice effect in all groups, and the large coaching effect among Africans—a total rise of over 14 points. The small coaching gain among the Indians was difficult to explain, but might be due to their anxiety to do well at first and their later lack of interest in non-verbal materials. Silvey (1964) also found a large practice rise, averaging 4·84 points, on the Progressive Matrices test in Uganda, even among selected secondary pupils, and a further 2·43 points after giving additional instruction.

TABLE 1

PRACTICE AND COACHING EFFECTS IN SOUTH AFRICAN SCHOOLS.
(Lloyd and Pidgeon).

	European	African	Indian
Mean initial I.Q.	103·2	86·8	86·8
Practice gain	7·39	6·95	5·65
Additional coaching gain	3·21	7·60	0·45

II.—THE PRESENT INVESTIGATION.

In 1965, the writer had been invited to survey and advise on problems of secondary school selection in Tanzania by the Ministry of Education (Vernon, 1966a), and on returning to work at University College, Dar es Salaam, in 1966, he was permitted to carry out an investigation in Tanzanian primary schools. Moreover, the Ministry, kindly made available the assistance of an inspector to act as tester in Swahili.* The object of the research was not to devise tests for secondary school selection in African countries, but to help to clear the ground by studying the effects of different techniques of administration.

Three types of nonverbal subtest were chosen, Number Series, Classification of Figures or Odd-Man-Out, and Matrices with creative responses, i.e., the pupil drew in his own solution. Three parallel forms of each test, containing twenty to twenty-five items, were constructed in England, tried out and item-analysed on six classes (Junior, 2, 3 and 4) in a large primary school. Full instructions, such as might be used if the test was applied without any oral explanation, were duplicated in English and in Swahili. The time limits found suitable in Tanzania, without involving speed stress, were twenty, ten and twelve minutes, respectively.

The general plan of the experiment was to give one test to each of several Standard 7 (top primary), classes, with printed instructions in English or Swahili, or with minimal oral instructions (the printed version being read out). A few days later each class was retested with a parallel form, either under the same conditions, or with the instructions in a different language, or with fuller explanation or coaching. Fourteen classes in six schools in Dar es Salaam took part, and a total of 554 pupils completed both tests. The three forms of the test were rotated so as to arrive at equivalent scores, and all results were converted to standard scores with a mean of 100 (for previously tested groups) and S.D. 15. No account was taken of age. Just one third of the pupils were girls.

Five conditions of administration, and two languages, were compared:

P=Printed instructions only. E=English.
O=Same instructions read orally. S=Swahili.

B=Full oral explanation with blackboard demonstration of additional items from the third form of the test. Pupils were asked to suggest and explain answers. Up to a quarter-of-an-hour was spent on each test, and this procedure should be comparable to that of Schwarz.

C=Coaching. Each pupil was given back his own previous test, already scored. Some five items were selected in each test, which many pupils had answered erroneously, and these were explained, as in B.

I=Individual coaching. Two pupils were taken at a time by each tester. The first was given his previous, marked, script; his wrong answers to five items in Subtest 1 were explained He then started on the parallel form and the second pupil was coached; similarly, for Subtests 2 and 3. The coaching had to be limited to five minutes per pupil per subtest, but it could be conducted at the pupil's own level.

---

* I am most grateful to the Chief Education Officer of the Ministry for releasing Mr. Mwalongo, who had been trained in testing techniques at E.T.S., Princeton, N.J. I would thank him for his willing co-operation and meticulous adherence to the requirements of the experimental design. The Ministry and the six head teachers were most helpful in providing facilities. Travel and other expenses of the research were met by a grant from the Association for the Aid of Crippled Children (New York).

The following fourteen comparisons appeared to be the most informative (PE=Printed English, OS=Oral Swahili, etc.) :

Group	First Test	Second Test
1.	PE	PE
2.	PS	PS
3.	OE	OE
4.	OS	OS
5.	PE	PS
6.	OE	OS
7.	PE	OE
8.	PS	OS
9.	OE	BE
10.	OE	CE
11.	OE	IE
12.	OS	BS
13.	OS	CS
14.	OS	IS

Ideally, all pupils should be assigned randomly to these fourteen groups ; but this would be quite impracticable in schools which have no classrooms to spare. Occasionally, two classes were mixed and reallocated for second testing but for the most part each group was represented by a single class. Most classes (apart from those in one school, mainly populated by pupils of Indian race) were probably very similar in ability, but this was an admitted weakness of design. Another weakness was that it was not possible, at this time of year, to find local students or teachers to do the testing in English, hence this was done by the writer or his wife, while the testing in Swahili was done by the Ministry inspector, Mr. Mwalongo. Thus, to some extent school differences and race-of-tester differences may be confounded with test-administration differences.

Co-operation was satisfactory, but it was difficult to gauge how well motivated the pupils were, especially when faced with complicated printed or oral English instructions. A major trouble was to get them to keep to one subtest at a time. Many turned on or turned back, and though this was checked it may explain some of the extreme score changes. On the other hand, there was virtually no trouble with copying from neighbours, except in the school consisting mostly of Asian pupils. These Indians (in contrast to Lloyd's observation) appeared more highly motivated throughout, but they were inveterate copiers ; and several scripts were later discarded when scoring revealed identical patterns of errors between neighbours.

III.—RESULTS.

(i) *Initial Means.* The initial means for five sets of classes are shown in Table 2. Although, as just mentioned, these may be affected by school differences, they indicate that English instructions severely increase the difficulty of

TABLE 2

INITIAL MEAN STANDARD SCORES.

No. of Classes	Condition of testing	Number Series	Classification	Matrices	Average
3	PE	93	92	99	95
2	PS	96	98	95	96
3	OE	94	97	99	97
4	OS	100	99	97	99
2 Asian Classes	OE	101	100	102	101

Number Series and Classification, though not Matrices which is almost self-explanatory from its sample items. All tests are somewhat affected also by giving the instructions in printed form only, instead of orally.

(ii) *Gains.* The gains for each of the fourteen groups are listed in the Appendix, together with their S.Es. It will be seen that these are highly variable from group to group and from one subtest to another, ranging from 1 point of Standard Score to 13. Individual pupils' raw score changes ranged from + 11 to –9. Nevertheless, the correlations between first and second tests were mostly in the ·70's or ·80's, occasionally sinking to 0·50, which is not unduly low considering the shortness of the subtest, and the intervention of coaching, discussion among the pupils, or other changes before the second testing. This implies that few comparisons can safely be made between particular pairs of classes. We must seek greater reliability by combining classes into groups, as has been done in Table 3.

TABLE 3

MEAN GAINS AND STANDARD ERRORS IN FIVE MAIN GROUPS.

Groups	Treatments	N	No. Series	Classification	Matrices	Average
1, 5	PE-PE, PE-PS	70	7·34±1·05	8·78±2·12	5·66±1·54	7·26±1·57
3, 7	PE-OE, OE-OE	91	3·41±0·91	3·26±1·50	3·66±1·08	3·44±1·16
2, 4, 6, 8	PE-PS, OS-OS, OE-OS, PS-OS	151	3·14±0·70	4·43±1·30	5·50±0·79	4·33±0·93
9, 10, 11	OE-BE, OE-CE, OE-IE	136	6·98±0·74	6·15±1·23	4·04±0·91	5·73+0·96
12, 13, 14	OS-BS, OS-CS, OS-IS	106	6·71±0·80	7·86±1·22	6·56±1·01	7·04±1·01

Line 1 shows the combined results for Groups 1 and 5 (PE–PE and PE–PS) which appear quite discrepant from the general trend of the remainder. Here the gains are very much higher than would be expected, in view of the absence of coaching, averaging 5·19 and 9·33 I.Q. points in the three subtests. By contrast Group 7 (PE–OE) in the same school, showed a much smaller gain of 2·60 despite the advantage of the switch to oral instructions on the second occasion. Note also that five of the six S.Es. of gains in Groups 1 and 5 are the highest in the whole list, i.e., there was more variability here than in any other classes. Almost certainly the reason must be that, since it was feared that the groups receiving printed instructions only might fail to maintain sufficient interest and concentration at the second test, they were given a special 'pep talk' by Mr. Mwalongo; and he seems to have been all too effective, Group 7 was not given such a talk. We must infer that the performance on intelligence tests of unsophisticated pupils such as Africans is more susceptible to motivational or attitude effects than is usually found in American or British researches; and to some extent this invalidates the rest of the research since we have no means of knowing whether other uncontrolled attitudinal factors did not affect the work in other classes. Nevertheless, the grouped results of the other twelve classes show a much more logical pattern, and the S.Es. of mean gains are now quite small.

Line 2 shows the mean gain from practice with oral English to average only 3·4 points, while practice either on printed or orally administered tests in Swahili (Line 3) is slightly higher at 4·3. Additional explanations or coaching by

English testers produces a substantial rise to 5·7 points, and by the Swahili tester to 7·0. Note that the ratio of gains for Swahili over English is almost identical under the practice and the coaching conditions. However, the additional gain from Swahili clearly does not affect Number Series at all, presumably because number teaching in schools is already largely given in English, whereas it is more difficult for pupils to understand what to do with totally unfamiliar non-verbal materials when instructed in English.

Perhaps the most striking feature is that the figures are not larger. They are more comparable to the results of practice and coaching experiments in England than to those quoted by Lloyd and Pidgeon. True, the latter gave more extended instruction between their first and final tests, but it is well established that additional experience beyond a single practice test and half-an-hour or so of coaching produces relatively little additional gain. The probable reason is that these Tanzanian pupils had recently taken their '11+' or General Entrance Examination, and had been trained to answer multiple-choice or short-answer achievement items, whereas the South African groups would have been less sophisticated to tests of any kind. One may presume that they were able to transfer their familiarity with printed tests to new intelligence test materials.

Comparing Table 3 with the more detailed figures in the Appendix, it is obvious that there are variations among the classes in each of our five groups, though they may be largely attributable to chance. But it is interesting to observe that there are no consistent differences between the three methods of explanation or coaching. The English testers do much better with blackboard instruction, worst with individual, the Swahili tester the reverse. It may well be that some African pupils were put off by being taken out of class by white testers. Also, one advantage of the group blackboard method was that the tester often called on brighter children to explain items to the rest of the class in their own English. Note, however, that the reliability of gains is no lower for the English than the Swahili tester. Another point of interest is that there is little indication of lower gains in the predominantly Asian school (Groups 3 and 10) than in the all-African, such as Lloyd found.

(iii) *Gains at Different Ability Levels.* A clearer picture of the differential effects of practice in Groups 2–4 and 6–8, and coaching in Groups 9–14, may be obtained by separating off the initially high and low scorers on each subtest. Normally, in accordance with regression effects, one would expect those above average to drop at second testing, those below average to show the biggest gains (cf. Wiseman and Wrigley, 1953) these differences being more marked the lower the correlation between the two testings, or the greater the S.E. of gains. Arbitrary dividing lines were set cutting off approximately the top 7 per cent. and the bottom 22 per cent. on each initial subtest. Table 4 shows the results.

TABLE 4

GAINS AFTER PRACTICE AND COACHING IN HIGH AND LOW GROUPS.

	After Practice in Groups 2-4 and 6-8			After Coaching in Groups 9-14		
	No. Ser.	Class.	Matr.	No. Ser.	Class.	Matr.
Top scorers . .	−2·65	−10·30	−2·41	+2·50	−5·34	−2·30
All cases	+3·28	+3·73	+4·67	+7·06	+6·75	+5·29
Low scorers	+4·61	+13·80	+5·68	+7·14	+15·70	+11·01

Number Series shows a moderate overall gain with practice (3·3 I.Q. points), much more with coaching (7·1). But the figures for the low scorers are very little larger. They fail to regress to the mean, showing that neither practice nor coaching helps them much. On the other hand, the top scorers drop very little with practice and actually gain after coaching; they definitely benefit from having items explained.

Classification, being the least reliable test, shows the biggest gains among the 'bottoms' and losses among the 'tops.' Coaching is more effective than practice at every level. On the Matrices test, coaching has scarcely any greater effect than practice, except among the 'bottoms.' Average and brighter children apparently do not benefit from explanation, whereas poor scorers do improve considerably more than from practice alone.

(iv) *Correlations with School Achievement.* It was hypothesized that the correlations between test scores and teachers' judgments of ability would rise at the second testing, when pupils were more familiar with what they had to do. Thus, correlations were calculated between initial or final tests and the overall school grade for the year. These grades are likely to be biased by achievement in formal subjects and by character qualities, at least in some schools; hence, high correlations should not be expected, especially with non-verbal tests. Both the grades and test results were split close to the 67th and 33rd percentiles; two values of tetrachoric r were obtained and averaged. The coefficients appear in the last column of the Appendix.

It will be seen that they vary considerably from one class to another, and that only eight of the fourteen show rises at the second testing. The overall means are 0·648 and 0·640 for first and second testing. However, on classifying according to the nature of the second test we find that:

In four groups given C or I (group or individual coaching) there are decreases. The mean r's are ·77 and ·61.

In seven groups taking the second test in Swahili, or given blackboard instruction, there are rises. The means are ·60 and ·67.

In three groups taking the second test in English, the results are irregular; means ·59 and ·62.

Thus, the hypothesis receives some confirmation when the instructions for the second test are in the easier language (regardless of the language of the first test); but it may be upset when the instructions are in English, or when detailed explanations of previous errors are given in either language. In some way, which we cannot explain, individual or group coaching seems to distort the scores and produce rather poorer agreement with school gradings.

It is interesting that the correlations were highest in the school which appeared to us the most lively, and where the Standard 7 classes are often taken by an American Peace Corps teacher. The gradings of achievement in this school may well have taken more account of the pupils' intelligent thinking than in the others. In Group 9, the correlation of ·95 means that thirty-three out of forty-two pupils were identically placed in top, middle and bottom thirds by the school and by the test. The other nine were misplaced by one grade only.

(v) *Item Difficulties.* The investigation had a subsidiary object, namely to test Irvine's (1966) hypothesis that, when a test measures different abilities in different cultures, the order of item-difficulties will alter. Some items which are relatively easy for English pupils will be inappropriate for Africans and will be answered relatively badly, others the reverse. The present writer expected the order of difficulty of Number Series items to be unaffected in so far as the order of acquisition of number relationships is likely to be pretty universal; whereas it would not be surprising if shapes, diagrams and their relations were perceived differently by children in different cultures.

Item pass rates were tabulated for the three versions of each subtest in three groups of approximately eighty pupils: (i) Tanzanian pupils at first attempt; (ii) Tanzanian pupils after practice or coaching; (iii) English pupils. The mean rank correlations were:

	Number Series	Classification	Matrices
(i) with (ii)	·936	·914	·957
(iii) with (i) or (ii)	·952	·884	·889

The hypothesis is, therefore, confirmed, for though the orders of difficulty for the two nonverbal tests are still very similar between English and Tanzanian (over 0·88), yet they are lower than when Tanzanian groups are compared. An attempt was made to analyse what sorts of items showed the greatest discrepancies in order of difficulty, but no generalization seemed possible. Thus there was no indication that the African children are relatively more successful at items which can be perceived 'holistically,' or do worse than English on items requiring analysis of details, though this view has been proposed by Biesheuvel (1966). An introspective analysis of the way Africans solve, or fail to solve, the specially difficult items might be illuminating.

IV.—DISCUSSION AND CONCLUSIONS.

Taking the last point first, it would follow that test items intended for pupils in a culture very different from that of Western Europe and North America should preferably be formulated by psychologists who are members of that culture, in order to avoid ideas and relationships which are inappropriate to the language and modes of perception and thinking of the culture. This may be unnecessary in the case of items using a numerical medium, but might be even more important for pictorial or verbal items than for abstract shape items, since they might arouse very different associations from those aroused among Western children. Clearly also such items should be analysed for difficulty and consistency in a sample of the population for whom the test is intended.

The main finding of the study is that practice and coaching effects on subtests in numerical and visual media, among African pupils who have never seen intelligence tests before, are no larger than 3 to 8 Standard Score units, apart from variations which could be attributable to chance. If British experience is any guide, the effects on verbal materials would be less than on nonverbal. The reason suggested is that these pupils are already practised in answering multiple-choice and short-answer scholastic achievement tests. This implies that similar tests based on 'finding the sequence,' 'odd man out,' or multiple-choice formats could be used among Tanzanian pupils at the end of primary schooling with only a small amount of previous familiarization, similar to that often provided in England. For example, a sample test could be released for practice purposes to all schools shortly before the secondary school selection examinations, and the teachers given guidance on explaining the various kinds of items.

It is true that the experiment has not demonstrated that more intensive coaching might not produce greater gains, which might be very uneven from one school to another. But the absence of significant and consistent differences between our various methods of coaching—the most intensive individual method being less effective with English testers than group instruction at the blackboard—suggests that, as in England, most of the possible gain is brought about by one practice test and a small amount of further class instruction.

The results with two classes given a pep talk indicate that motivation and set factors can play a larger part in performance than differences in practice and coaching. Admittedly, this finding seriously reduces the value of the experiment as such. But under conditions of maximum motivation in all schools, as in selection examinations, test reliability should not be impaired.

In view of the superiority of oral over printed instructions, and Swahili over English, one would like to recommend that Swahili instructions be given orally in Tanzania, or other countries with similar linguistic difficulties. But the selection examinations are applied in different schools by some 1,500 invigilators, who are teachers, not trained testers. If they were allowed to read the instructions, difference in manner, clarity or elaboration of instructions might be considerable. Hence, it is preferable to print them in Swahili (or whatever language has been most extensively used for school instruction), and to let the pupils' own teachers read and explain them to poor readers during the practice testing.

The experiment was not designed to show what kinds of test materials are most reliable and valid. The three subtests used here have worked well, though initially good and poor scorers are affected by practice and coaching in an uneven manner. In the average school the correlations between the total initial or final test and the school's own achievement at gradings were reasonably high, and they showed slight improvements at the second testing after practice or blackboard instruction, though not after other types of coaching. However when tests are used for predicting aptitude for secondary work, it would certainly be necessary to rely more on verbal items, whether in English or Swahili. In view of our difficulties in preventing pupils from wandering from page to page, there would be a clear advantage in using the cycle-omnibus type of test (like Moray House and N.F.E.R.). If Matrices items were included, they would have to be multiple-choice for ease of scoring.

The next stage in the introduction of intelligence tests should consist of trials of a wide variety of item-types and their validation. Each type should be judged on several criteria :

(i) Ease in getting across with printed instructions only.

(ii) Small gains with practice and coaching.

(iii) Small differences between the scores in efficient schools with well-qualified staff and smaller schools in remote areas, which may suffer from greater linguistic handicaps.

(iv) Suitable distribution of item-difficulties, and acceptable reliability.

(v) Correlation with teachers' ratings of intelligence, carefully defined to distinguish it from formal achievement and character traits.

V.—REFERENCES.

BIESHEUVEL, S. (1949). Psychological tests and their application to non-European peoples. London : Evans Bros., *Yearbook of Education*, 87-126.

BIESHEUVEL, S. (1966). Some African acculturation problems, with special reference to perceptual and psychomotor skills. Unpublished paper, Wenner-Green Foundation for Anthropological Research, Symposium No. 33.

IRVINE, S. H. (1965). *Selection for Secondary Education in Southern Rhodesia.* University College of Rhodesia and Nyasaland, Faculty of Education, Occasional Paper No. 4.

IRVINE, S. H. (1966). Towards a rationale for testing attainments and abilities in Africa. *Brit. J. Educ. Psychol.*, 36, 24-32.

LLOYD, F., and PIDGEON, D. A. (1961). An investigation into the effects of coaching on non-verbal test material within European, Indian and African children. *Brit. J. Educ. Psychol.*, 31, 145-151.

OLIVER, R. A. C. (1932). Intelligence of European and African schoolboys in Kenya. *East African Med. J.*, 9, 160.

SCOTT, G. C. (1950). Measuring Sudanese intelligence. *Brit. J. Educ. Psychol.*, 20, 43-54.

SCHWARZ, P. A. (1961). *Aptitude Tests for Use in Developing Nations.* Pittsburgh: American Institute for Research.

SILVEY, J. (1964). *Selection for Secondary Schools in Uganda.* Kampala, Uganda: East African Institute of Social Research.

VERNON, P. E. (1966a). *Selection to Secondary Schools in Tanzania.* Unpublished Report to Tanzanian Ministry of Education.

VERNON, P. E. (1966b). Development of current ideas about intelligence tests. In *Genetic and Environmental Factors in Human Ability* (Edit. J. E. Meade and A. S. Parkes). Edinburgh: Oliver and Boyd, 3-14.

WISEMAN, S., and WRIGLEY, J. (1953). The comparative effects of coaching and practice on the results of verbal intelligence tests. *Brit. J. Psychol.*, 44, 83-94.

(*Manuscript received 23rd January, 1967*)

VI.—APPENDIX.

Mean Gains and S.Es. in each Class. Also Correlations with School Achievement

Group	School	N	Treatment	Gains and S.Es. (below)				Correlations
				No. Ser.	Class.	Matr.	Average	
1	A	30	PE-PE	7·63	4·72	3·23	5·19	·58
				±1·89	±3·02	±2·18		·40
2	B	33	PS-PS	3·14	2·13	5·80	3·69	·66
				1·75	2·42	1·62		·68
3	C	49	OE-OE	1·73	5·67	5·45	4·28	·36
				1·11	2·01	1·61		·62
4	D	39	OS-OS	3·45	7·04	6·42	5·62	·64
				1·18	2·51	1·51		·66
5	A	40	PE-PS	7·05	12·85	8·09	9·33	·56
				1·23	2·92	2·15		·65
6	B	41	OE-OS	1·24	3·31	3·11	2·55	·52
				1·40	2·57	1·54		·56
7	A	42	PE-OE	5·09	0·85	1·87	2·60	·83
				1·46	2·25	1·40		·83
8	B	38	PS-OS	4·72	5·24	6·46	5·47	·54
				1·32	2·89	1·66		·59
9	E	42	OE-BE	9·47	10·50	4·08	8·02	·87
				1·48	2·14	1·19		·95
10	C	53	OE-CE	5·52	4·02	4·71	4·75	·80
				1·13	1·80	1·52		·66
11	E	41	OE-IE	5·96	3·92	3·34	4·41	·82
				1·28	2·52	1·82		·74
12	F	42	OS-BS	8·08	6·66	2·88	5·87	·40
				1·33	1·42	1·36		·58
13	D	35	OS-CS	3·37	6·62	10·59	6·86	·69
				1·44	2·68	2·09		·48
14	F	29	OS-IS	8·67	10·30	6·22	8·40	·79
				1·36	2·29	1·93		·56

THE PERSONALITY PATTERN OF CREATIVE ARTISTS

By
PETER G. CROSS
(*The Firs Junior School, Castle Bromwich, Birmingham*)
RAYMOND B. CATTELL
(*University of Illinois*)
AND
H. J. BUTCHER
(*University of Edinburgh*)

SUMMARY. Sixty-three visual artists and twenty-eight craft students were compared with a matched control group. Significant differences in mean scores between artists and controls were found on twelve factors of the 16PF test. On eleven of these twelve factors, the scores of the craft students were intermediate between those of the artists and the controls. Especially salient features of the artists' personality pattern were A – (reserved, schizothyme tendency), E+ (assertiveness, dominance), Q2+ (self sufficiency), C – (low emotional stability), Q3 – (low self-integration, casualness), M+ (autistic or bohemian tendency) and G – (low superego strength). They also differed from the control group in being more suspicious (L+), more apprehensive or guilt-prone (O+), and more tense or overwrought (Q4+). In terms of the second order factors, the artists are assessed as being slightly introverted though there is evidence of some contradiction, strongly anxious, experimenting, non-moralistic, and slightly sensitive.

I.—RESEARCH ON CREATIVITY.

SINCE about 1950, the volume of research in the large and rather amorphous field of creativity has expanded almost exponentially, particularly in the U.S.A. Most of this work has been concentrated on the cognitive aspects of creative ability. In particular, the work of Getzels and Jackson (1962), in which it was claimed that highly intelligent and highly creative children form rather separate classes, has aroused considerable controversy. Attempts to replicate their findings in groups of children unselected for ability have met with mixed success (Torrance, 1962; Edwards and Tyler, 1965; Hasan and Butcher, 1966). Quite recently, further success in isolating creative abilities from general intelligence has been claimed by Wallach and Kogan (1965).

Interesting and important as some of this work is, it appears likely that non-cognitive factors are of at least equal importance with cognitive in distinguishing highly creative from less creative adults. Quantitative and non-impressionistic research in this area has been rare, but Cattell and Drevdahl, (1955), Drevdahl (1956), Drevdahl and Cattell (1958), Barron (1957, 1963) have established some clearly differentiating factors. In general, the same factors have distinguished both creative writers and eminent research scientists, for instance, both from the general population and from less creative men and women of equal intellectual ability.

The commonest findings in terms of factors measured by the 16PF test are that creative adults tend to be schizothyme (A –), dominant (E+), self-sufficient (Q2+), desurgent (F –), bohemian (M+), and radical (Q1+). In terms of the main second-order factors, introversion/extraversion and anxiety, the picture is quite complex. As regards the first, Cattell has characterised the typical pattern as " a constitutional endowment towards extraversion, modified by environmental influences into an introvert pattern." Barron (1963), interpreting results obtained on the MMPI with eminent writers, found them " both sicker and healthier psychologically than people in general. Or, to put it another way, they are much troubled psychologically, but they also have far greater resources to deal with their troubles." Some of these findings will be reviewed in more detail in a later section, where comparisons with the results of the present study are made.

II.—METHOD.

(a) *Selection of the samples.*

The sixty-three artists (thirty-six male, twenty-seven female) were selected as having given clear evidence of unusual talent in drawing and painting. Fourteen were professional art teachers or professional painters, seven were students at the City of Birmingham Training College (selected from about 120 taking the Art and Craft course), twelve were students at Birmingham Art College (selected from about sixty taking the N.D.D. course in painting), five were students at the St. Martin's Art College, London (selected from about forty taking the pre-diploma course), twenty-five were students studying for the Art Teacher's Diploma at Birmingham Art Teachers' Training College. All the students in the sample were considered to be especially gifted in the art of drawing and painting, and were rated as entirely capable of taking up a career in art.

The twenty-eight craft students were students at the Art Teachers' Training College who were also studying for their Art Teacher's Diploma, but who had successfully specialised in various different crafts at Art College. They include, for instance, seven potters, six dress designers, one silversmith, and one furniture designer.

The sixty-three controls (thirty-six male, twenty-seven female) consisted of lecturers and students at the City of Birmingham Training College, none of whom practised painting or drawing even as a hobby. About half were science specialists, and half specialists in 'arts' subjects such as English, History, Geography, Scripture. The number of lecturers matched the number of art teachers, lecturers and professional artists in the artist sample. The two groups (artists and controls) were also carefully matched for age (males and females separately) and for educational background.

Mean age of both artists and controls was 24, of craft students 22.

(b) *Administration of the* 16*PF Test.*

Forms A and B were administered individually, or in small groups. Subjects were allowed to remain anonymous if they wished.

III.—RESULTS.

(a) *Comparison of the three groups.*

Table 1 shows the means and standard deviations on the sixteen factors for the three groups. The last three columns show the results of tests of significance between the group means.

Of fifteen personality factors (excluding factor B, intelligence), twelve discriminate significantly between artists and controls, including ten at the ·01 level of significance. These ten are A −, E+, Q2+, C −, L+, O+, Q3 −, Q4+, M+ and G − (in each case, the + or − sign indicates the 'artist' pole). These differences are interpreted as showing the artists to be more withdrawn,

TABLE 1

COMPARISON OF ARTISTS, CONTROLS AND CRAFT STUDENTS ON SIXTEEN PERSONALITY FACTORS.

Factor	1 Artists (N—63)		2 Controls (N—63)		3 Craft stds. (N—28)		Signif. of diffs.		
	Mean	S.D.	Mean	S.D.	Mean	S.D.	1 *v* 2	1 *v* 3	2 *v* 3
A (Schizothymic tendency ' sizothymia ')	16·30	5·20	19·35	4·62	17·89	5·37	·001		
E (Dominance, assertiveness)	28·59	6·86	23·65	6·50	24·75	5·86	·0001		
F (Surgency)	29·37	7·66	29·29	9·50	29·79	6·96			
H (Social boldness)	22·16	9·36	22·86	9·66	24·18	9·62			
Q2 (Self sufficiency)	26·41	4·69	22·29	4·95	23·18	5·00	·0001		
C (Ego strength, emotional stability)	23·73	7·32	29·24	6·66	27·64	7·88	·0001		
L (Suspiciousness, ' protension ')	20·21	5·10	17·56	5·19	18·18	4·93	·01		
O (Apprehensiveness, guilt proneness)	27·95	7·75	24·57	6·76	23·07	7·91	·01	·05	
Q3 (Self control, social self discipline)	17·02	4·51	20·83	5·77	19·32	6·29	·0001		
Q4 (Ergic tension)	31·71	8·26	27·57	8·30	28·21	8·74	·01		
I (Tender-mindedness, sensitivity, ' premsia ')	24·30	4·52	22·19	4·97	23·11	4·50	·05		
N (Shrewdness)	18·29	4·50	19·90	4·45	18·64	3·85			
M (Autistic or bohemnian tendency, ' hysterical unconcern ')	33·37	5·34	25·25	7·24	30·79	5·80	·0001		·001
Q1 (Radicalism)	23·37	4·98	21·63	4·58	21·86	4·02	·05		
G (Super ego strength, conscientiousness)	17·71	5·44	24·22	6·84	20·18	6·32	·0001		·01
B (Intelligence)	19·48	2·49	19·52	7·70	18·86	1·69			

or schizothyme, of higher dominance or assertiveness, of lower emotional stability or ego strength, of higher suspiciousness or self-projection (called protension), of greater apprehensiveness or guilt proneness, of greater casualness or lower self-integration, of higher ergic tension, of higher autistic tendency or ' hysterical unconcern ' and finally of lower superego strength.

Only two differences between craft students and controls are statistically significant. The craft students (like the artists, but to a lesser degree) differ from the controls in respect of lower superego strength and higher ' autistic ' tendency. The smaller number in the craft student group will tend to preclude statistically significant differences, but clearly a major feature of the results is that the craft students' scores are, in fact, intermediate between those of artists and controls on almost every factor. Inspection of the table of means reveals a remarkable regularity. On every factor which discriminates between artists and controls, except for 0+ (apprehensiveness) and Q4 (ergic tension), the craft students' scores are intermediate, but usually closer to those of the controls.

(*b*) *Comparison of the three groups with student and general population norms.*

Although the 16PF has been used fairly extensively in Britain, adequate national norms are not yet available. Thus, the comparisons shown in Table 2 are first with American general population norms and second, with American student norms, since a large proportion of artists, craftsmen and controls in

TABLE 2

Sten Score (Standard Ten Score)—Comparison of Artists, Controls, Craft Students and Research Scientists, Using American Population Norms.

		Minimum 1 Maximum 10 Average Sten Scores (5·5 mid-point)			
16 PF Factor 2nd Order Factor		Scientists N—144 *GP	Artists N—63 GP*(SP)	Craft Stds. N—28 GP*(SP)	Controls N—63 GP*(SP)
I—General Extraversion :					
Cyclothymia (outgoing, participating)	A	3·4	4·3 (4·2)	4·6 (4·4)	5·1(5·2)
Dominance (assertive, aggressive)	E	7·2	7·7 (7·5)	6·5 (6·4)	6·2 (6·0)
Surgency (enthusiastic, gay)	F	3·5	6·2 (5·6)	6·3 (5·4)	6·2 (5·4)
Parmia (socially bold, venturesome)	H	6·5	4·5 (4·6)	4·9 (5·2)	4·7 (4·8)
**Group adherence (lacking in self-sufficiency)	11–Q2	4·5	3·2 (3·2)	4·3 (4·1)	4·7 (4·4)
II—General Anxiety or Neuroticism:					
**Lower Ego Strength (emotionally less stable)	11–C	4·1	8·0 (7·6)	6·9 (6·3)	6·5 (6·1)
Protension (suspicious, self-opinionated)	L	4·1	7·2 (6·9)	6·6 (6·3)	6·2 (6·0)
Guilt Proneness (apprehensive, worried)	O	3·6	7·4 (7·2)	6·1 (5·7)	6·5 (6·2)
**Low Integration (lacking in self-control)	11–Q3	4·2	8·0 (7·4)	7·0 (6·3)	6·4 (5 9)
Ergic Tension (tense, driven, overwrought)	Q4	5·1	7·5 (7·0)	6·7 (5·8)	6·5 (5·1)
III—Sensitivity. :					
Premsia (tender-minded, sensitive)	I	7·1	6·7 (6·7)	6·2 (5·9)	5·9 (5·9)
**Naivete (lacking in shrewdness)	11–N	5·5	7·1 (6·7)	7·0 (6·2)	6·4 (5·9)
IV—Experimentation ·					
Autia (bohemian tendency, imaginative)	M	5·6	8·8 (8·8)	8·1 (8 0)	6·4 (6·3)
Radicalism (experimenting, free thinking)	Q1	6·2	7·4 (7·5)	6·8 (7·2)	6·7 (6·8)
V—Moral Concern :					
Superego strength (conscientious, rule-bound)	G	3·4	1·5 (2·7)	2·4 (3·7)	3·9 (4·8)
VI.—Intelligence :					
Higher Scholastic Mental Capacity (bright)	B	6·8	9·2 (7·7)	8·9 (7·2)	9·2 (7·7)

* GP—Average sten score, using general population norms.
SP—Average sten score, using student norms.

**Denotes factors where the sten score has been reversed to give the correct pole for its second order factor.

this study are students. Although the ' selection level ' for American students is lower (approximately one in four Americans have college education) than for the students in this research, it is clear that the student norms have the effect of reducing the extent of divergence from the 5·5 sten mid point for all three groups, particularly in the case of G – (lower superego). However, it is also noticeable that on both American norms the control group in the present study obtains a mean sten within one sten above or below 5·5 on thirteen out of fifteen factors (again excluding factor B, intelligence). In the case of factors Q2, C, Q3 and N, the sten scores are subtracted from 11·0 to give a correct alignment for the second order factors.

Table 2 outlines in some detail the salient features of the artists' mean personality profile. In terms of deviations from the American population norms (in units based on standard deviation), the outstanding characteristics in order of extent of deviation are M+, G –, C –, E+, Q2+, Q3 –. Thus, in the case of these six factors, one can say that the artists differ markedly both from the controls and from the test norms. The other four factors on which the artists clearly differ from the controls are A –, L+, O+, and Q4+. On A – and L+, the deviation from population norms is not so strongly marked. On O+ and Q4+, the deviation is stronger, but the controls also had higher sten scores on these factors. Considering the careful matching of controls, it may be said that I+ (emotional sensitivity), Q1+ (free-thinking radicalism), and N – (naivete), are also suggested by Table 2 as being characteristics of the artist group, though to a barely significant degree (see Table 1).

In terms of the second order factors, the artists show slight overall introversion (A –, H – and Q2+) though there is evidence of contradiction, with clear E+ and absence of indication on factor F. General anxiety or neuroticism is very strong (C –, L+, O+, Q3 –, Q4+) and so is experimentation (M+, Q1+) and moral disregard (G –). There is also slight, but consistent sensitivity (I+, N –). There was little difference between the scores of the three groups on factor B, intelligence, but clearly all three groups have an average intelligence well above normal The intermediate position on most factors of the craft students discussed in the preceding section is also obvious in Table 2. The main significance of this is that the craft students, therefore, seem to be generally less anxious than the artists. The artists' high anxiety level is discussed more fully in the following section.

(c) *Comparison between the characteristic personality patterns of Artists and Scientists.*

In Section I it was stated rather baldly that a common personality pattern was detectable among creative adults even though their fields of activity might be as different as creative writing and scientific research. It is interesting in this connection to compare the present results with those of other researches into scientific and artistic creativity.

Cattell and Drevdahl (1955) compared the personality patterns of eminent research scientists with those of other American academics on the hypothesis that the former patterns would be more characteristic of high creativity. The four factors on which highly significant differences were found ($p < {\cdot}001$ were A –, C –, M+, and Q2+ (the signs indicate the directions in which the research scientists differ from the administrators). The fact that these four factors are among the seven found to discriminate to the same significance level between artists and controls (in each case in the same direction), despite great differences in age, cultural background, eminence and field of work is highly suggestive of a common personality pattern among creative people. In another study on scientists (Jones and Cattell, 1966), involving industrial scientific work, M+ and Q2+ are again found to significantly differentiate creative individuals.

Drevdahl (1956) included separate Arts and Science groups in an investigation into creativity with a sample of graduate or advanced undergraduate students of Nebraska university. The arts group scored significantly higher than did the science group on M+ and I+, but the science group scored lower, though not significantly so, on A – and Q4 –. When the two groups were split into creative *versus* non-creative sub-groups, the creative arts students' scores were significantly higher than the creative scientists' on Q1 and Q2. Concerning the general group differences on M+ and I+, it could be said that there is a tendency for the Arts *versus* Science dichotomy to reflect tender-radicalism *versus* tough conservatism, a distinction that is well known politically, but not frequently found in personality characteristics in which toughness tends to be associated with radicalism rather than conservatism. There was an interesting interaction on factor L+, the creative arts and non-creative science students scoring high on this trait. The author concludes that suspiciousness and interest in the internal mental life may be associated with creativity in the arts, but that trusting attitudes and interest in external things rather than persons may be more associated with creativity in the sciences.

Cattell and Drevdahl (1958) found very large and consistent differences between creative writers and artists (N–153) and the general population on the 16 PF. In the anxiety primaries, they were high scoring on C+ and Q3+, but also on Q4+, i.e., they were basically stable people, living under strain. In extraversion, they were H+, F+ and E+, but also Q2+ and A –. On the remaining second order factors, their tendencies bear out the findings of this research, with clear findings on I+, M+, Q1+, and G –. It is interesting that they combine a strong self-sentiment and discipline with non-conformity to group superego standards. As Cattell points out, many of these are what Riesman classified as autonomous characteristics and creative persons are frequently not socially acceptable or popular. They are, in many ways, the exact opposite of the successful salesman. This suggests that, for purposes of academic selection, particularly for higher grade posts, extreme tolerance of what would otherwise be undesirable qualities is desirable. It is interesting that creative workers in both literary and artistic fields should differ in much the same way from less gifted persons. The main differences between the two groups was that the artists were even more tense, extraverted and radical than the writers.

The present study is, herefore, in considerable agreement with earlier work on temperamental aspects of creativity, particularly in its indication that artists tend to be high on M (perhaps the 'creativity' factor par excellence) and on Q2. In what respects does it add to or contradict pevious findings? Mainly, perhaps, in suggesting a higher anxiety level in the sample of painters than has been indicated in earlier studies of personality and creativity. This is particularly marked in contrast with studies of creative scientists. Of the ten chief characteristics shown by the artists, seven (C –, L+, O+, Q3 –, Q4+, M+, and G –) are known to load the second-order factor of anxiety. In the case of creative scientists, the evidence about anxiety level is complex, but in general it appears that there is no *overall* pattern of high anxiety. In the Cattell and Drevdahl (1955) study, for example, we have already mentioned that the research scientists, in marked contrast to our sample of artists, were significantly above the general population averages on factors C and Q3, and below the averages on L and O.

It is not really valid to make a direct comparison between artists' and scientists' scores, as the scientists' data were obtained using the 1957 version of the test, and American general population norms, and the artists' data, using the 1962 versions. However, this is rather more legitimate when both are converted to standard scores, as in Table 2. It is interesting to see that, although the scientists' and artists' scores both show clear A –, Q2+, E+, Q1+, I+ and G –, the largest factor differences are on the seven anxiety factors mentioned above, always in the direction of the artists' being more anxious. The only other large difference is that the scientists score lower on Factor F (surgency).

It is possible that the present sample of artists is not typical of all artists in this respect. It has been shown that Drevdahl and Cattell (1957) did not find general anxiety to be exceptionally high in their sample of artists and writers, which had high scores on C+ and Q3+. Unlike the present artist sample, few of the Drevdahl sample could have been students, being selected from the publication *Who's Who in American Art*. High general anxiety in students has been found in other studies. In his paper supporting the Maudsley Personality Inventory, Jenson (1958) reports giving the M.P.I. to various educational and clinical groups, most of them fairly large. In respect of neuroticism, high scores were found for both British and American university students, polytechnic and art students, and research workers, but not for industrial apprentices. Dysthymics, hysterics, and psychopaths also had high scores. In an article re-examining Jensen's data, Bronzaft, *et al.* (1960) claimed a negative correlation between extraversion and neuroticism and produced further data on female nurses to confirm their hypothesis. It is interesting that the artists of the present research exhibit, to some degree, the same correlation. In Britain, Warburton gave the 16PF to thirty-two students in a course of training for the teaching of art. The students' results showed anxiety (C –, O+, Q3 –) as well as introversion (H –, E –), unaffectedness (N –), radicalism (Q1+), non-conformity (M+) and lack of moral concern (G –).

Table 2 clearly shows that not only the artists but, to a lesser degree, the craft students and controls differed from the population norms in being more anxious. Thus, a possible but speculative explanation for a high level of anxiety in the present group would be that they are young adults, many about to leave or having shortly left the protection of an academic institution, and exposed to all the hazards of making a living in an artistic career.

IV.—REFERENCES.

BARRON, F. (1957). Originality in relation to personality and intellect. *J. Pers.*, 25, 730-742.

BARRON, F. (1963). *Creativity and Psychological Health.* New York : Van Nostrand.

BRONZAFT, A., HAYES, R., WELCH, L., and KOLTUV, M. (1960). Relationships between extraversios, neuroticism and ascendance. *J. of Psychol.*, 50, 279-285.

CATTELL, R. B., and DREVDAHL, J. E. (1955). A comparison of the personality profile (16 PF) of eminent researchers with that of eminent teachers and administrators, and of the general population. *Brit. J. Psychol.*, 46, 248-261.

CATTELL, R. B., and JONES, F. R. (1966). Prediction of high level creativity in industrial scientific work.

DREVDAHL, J. E. (1956). Factors of importance for creativity. *J. Clin. Psychol.*, 12, 21-26.

DREVDAHL, J. E., and CATTELL, R. B. (1958). Personality and creativity in artists and writers. *J. Clin. Psychol.*, 14, 107-111.

EDWARDS, E. P., and TYLER, L. E. (1965). Intelligence, creativity and achievement in a non-selective public junior high school. *J. Educ. Psychol.*, 56, 96-99.
GETZELS, J. W., and JACKSON, P. W. (1962). *Creativity and Intelligence.* New York : Wiley.
HASAN, P., and BUTCHER, H. J. (1966). Creativity and intelligence. A partial replication with Scottish children of Getzels and Jackson's study. *Brit. J. Psychol.*, 57, 129-136.
JENSEN, A. R. (1958). The Maudsley Personality Inventory. *Acta. Psychol.*, 14, 314-325.
TORRANCE, E. P. (1962). *Guiding Creative Talent.* Englewood Cliffs, N.J. : Prentice-Hall.
WALLACH, M. A., and KOGAN, N. (1965). *Modes of Thinking in Young Children.* New York : Holt, Rinehart and Winston.
WARBURTON, F. W. (1966). *The 16 PF Profile of a Group of Students on a Course of Training for Teachers of Art.* (Unpublished manuscript.)

(*Manuscript received 10th August,* 1966)

CORRELATION BETWEEN STREAMING AND SEASON OF BIRTH IN SECONDARY SCHOOLS

By P. SUTTON

(*Sarum, Fairview Drive, Hythe, Southampton*)*

SUMMARY. Data obtained from two grammar schools and six secondary modern schools in a fairly homogeneous area in Hampshire were analysed to discover the effects of season of birth on secondary school selection. It was found that a marked bias existed in stream selection in secondary modern schools, and that this remained constant up to school leaving No significant bias was found in 11 + selection.

I.—INTRODUCTION.

THERE is a great deal of published information on the relationship between season of birth and scholastic performance. The main conclusion which is now clearly established is that Autumn-born children, as a whole, show better performance than Summer-born. There is, however, some conflict of evidence and opinion in certain aspects of the general picture Armstrong (1966) states that " The performance of a large age group at 11 does not support the view that standardised tests discriminate against the Summer-born children." Jinks (1964) and Freyman (1965) found that 11+ selection was affected by season of birth. The explanation of this may be that 11+ selection is influenced to a varying degree under different authorities, by factors other than the results of standardised tests. On the subject of academic achievement, Armstrong states that there is no difference between Summer and Autumn-born grammar school children in G.C.E. results Freyman found that in secondary modern schools, there is a correlation between season of birth and stream position, and thus ultimate achievement.

II.—SURVEY.

A survey was made of secondary schools in a relatively homogeneous area in South Hampshire. The County Education Officer was unable to provide the data required, so it was necessary to request this from the individual schools. Three grammar and eight secondary modern schools were approached, and full data were obtained from two grammar and six secondary modern schools. There is no reason to suppose that the lack of response from three schools makes the results unrepresentative.

For the purposes of this study, data for the first three years only were used. This eliminates any bias resulting from early leavers, and other effects such as course streaming which are pronounced in the fourth year.

III.—STREAMING.

(*a*) *Secondary Modern Schools.*

All secondary modern schools streamed from the first year, but some re-arrangement was made in the fourth year, where classes were divided into classifications such as G.C.E., C.S.E., Commercial, Practical 1, Practical 2, etc.

---

* This work was carried out by the Eastleigh and Totton Associations for the Advancement of State Education, under the Chairmanship of the Author.

The majority of schools had four streams in most years ; there was only one with less than four streams in one of the first three years. Some years had more than four streams and one school had eight streams ; this last school used different titles, but the classes corresponded to regular streams. Where there were more than four streams, these have been grouped. For example, where there were five streams, D and E (or D and Remedial) are grouped together as D, and where there were eight streams, these are grouped first and second as A, third and fourth as B, etc. Table 1 shows the distribution of children into the various streams by birth groups.

TABLE 1

STREAM DISTRIBUTION BY BIRTH GROUPS.

Stream	A	B	C	D	Totals
(*a*) Sept.—Dec.	277	223	185	158	843
(*b*) Jan.—Apr.	230	234	221	180	865
(*c*) May—Aug.	207	205	221	252	885
Totals	714	662	627	590	2,593

$X^2 = 40{\cdot}5$ For 11 d.f., $p < {\cdot}001$

To correct for the slight inequality between birth groups, the number in each birth group in each stream was multiplied by the appropriate factor, the equalised numbers were then used in Table 2. The symmetry of the distribution of percentages is quite remarkable.

TABLE 2

STREAM PERCENTAGES BY BIRTH GROUPS (EQUALISED).

	A	B	C	D
(*a*) September—December	40	35	30	28
(*b*) January to April	32	35	35	30
(*c*) May—August	28	30	35	42
Totals	100	100	100	100

To investigate the consistency of this distribution through the schools, the percentages were calculated for the three years separately. Comparing the first and third years, it will be seen that for the twelve groups in each year, the percentage increased in five cases, decreased in four cases and remained constant in three cases. Thus, the distribution is consistent between the three years (Table 3).

TABLE 3

DISTRIBUTION BY STREAM AND BIRTH-GROUP THROUGH THE FIRST THREE YEARS.

(Percentage of total in each stream in each year)

Stream	A			B			C			D		
Birth Group	a	b	c	a	b	c	a	b	c	a	b	c
First Year	44	28	28	34	33	33	27	39	34	26	35	39
Second Year...... .	38	33	28	32	38	30	36	31	35	27	28	45
Third Year	37	35	28	37	33	30	29	37	34	30	28	42
Change: first year to third year	–7	+7	0	+3	0	–3	+2	–2	0	+4	–7	+3

(*b*) *Grammar schools.*

One of the grammar schools did not stream at all, and the other did not stream until the third year. To some extent, this represented course streaming. As a matter of interest, both third and fourth year streams were analysed (since it was considered that there would have been few, if any, leavers from the fourth year). The stream distribution by birth groups gave $\chi^2=15\cdot4$ ($\cdot10<p<\cdot20$) and showed no symmetry.

IV.—SELECTION AT 11+.

The survey did not give a complete sample of 11+ selection in any particular year. However, a comparison of populations of secondary modern and grammar pupils within a common catchment area and covering the same period of years will yield similar results. The available data covered 2,593 secondary modern and 590 grammar pupils. Thus, 18·5 per cent. of the total were in grammar schools, and this is very close to the average proportion selected at 11+ in the area (Table 4).

TABLE 4

SECONDARY MODERN AND GRAMMAR DISTRIBUTION BY BIRTHS GROUP

(1st, 2nd and 3rd years)

	Sept.-Dec.	Jan.-Apr.	May-Aug.	Total
Secondary Modern	843	865	885	2,593
Grammar	200	204	186	590
Total	1,043	1,069	1,071	3,183
Secondary Modern % ..	80·7	80·9	82·6	81·5
Grammar %......... . .	19·3	19·1	17·4	18·5

Taking the grammar school selection only, $\chi^2=1\cdot29$ For 2 d.f., $\cdot50<p<\cdot60$.

A decrease in percentage grammar school places from Autumn to Summer groups would be expected if there were a season of birth bias, but the value of χ^2 here is too low to allow any significance to be attached to it.

V.—CONCLUSIONS.

It is clear that streaming in secondary modern schools is biased by season of birth. This bias corresponds with that found by Jackson (1964) in junior schools, and presumably stems from the same cause. Unless this bias is overcome in the junior school it is likely to persist throughout a child's school career. On the other hand, it appears that 11+ selection based on standardised tests is relatively free from this bias.

It may well be that the child of high intelligence is better able to overcome this 'age group position' effect than the child of average or lower intelligence. If so, this would contribute to some extent towards the apparent success of 11+ selection, the absence of bias in grammar school results (Armstrong) and also the pronounced bias found among the E.S.N.

The implications of this to Comprehensive secondary education are serious. Streaming is deeply entrenched in the British educational system. Season of birth is only one factor which affects 'attainment' but not 'ability.' If standardised tests are eliminated, the effect of Comprehensive education could be to create new obstacles to equality of opportunity for many children. Stream selection must be examined in this context.

With regard to ultimate achievement, G.C.E. results are only one measure of achievement, and relate to only a minority of the school population. In secondary modern schools, it is clear that season of birth is related to achievement as measured by stream in the final year. One effect discovered in this survey was that the proportion of Spring and Summer-born children staying on beyond 15 in secondary modern schools was twice that of Autumn-born (because of the attraction of Christmas leaving for the Autumn-born). This affected the upper streams as well as the lower, and would tend to counter the stream selection bias towards the Autumn-born obtaining a higher proportion of examination passes.

The main conclusions drawn from this survey, taken in conjunction with other published material, are:

(1) Season of birth has a marked effect on ultimate achievement, at least insofar as secondary modern scholars are concerned (i.e., the majority of the population).

(2) The bias in streaming in secondary modern schools remains uniform up to school leaving.

(3) 11+ selection is comparatively successful in overcoming this bias.

(4) Because of the success of 11+ selection, and possibly other factors, this bias is not to be expected to any marked extent in G.C.E. results.

(5) Stream selection in both junior and secondary moderns chools is greatly affected by season of birth, which is virtually unrelated to innate ability. This sytem of selection is, therefore, inefficient and represents a serious obstacle to equality of opportunity.

ADDENDUM.

Since this manuscript was written, further research papers on this subject have been published, and the Plowden Report (which should, and generally does, reflect research findings) has appeared. The distribution shown in Table 2 corresponds to a difference in average age between top and bottom streams of 1·1 months. This distribution agrees fairly closely with that found by Jackson in four-stream junior schools (Jackson's data indicated a difference in average age of about 1·3 months in four-stream and 1·0 months in three-stream schools).

Lunn (1967) has found a difference of 2·3 months in four-stream and 1·5 months in three-stream junior schools. In addition, Lunn showed that the average age differences did not change significantly through the four years in junior school. This suggests that, as yet, junior schools are making no progress towards correction of the season of birth bias. A further important finding by Lunn is that only about 50 per cent. of junior schools employing homogeneous streaming use any form of standardised test to aid selection.

All this evidence indicates that stream selection is influenced primarily by immediate attainment rather than by innate ability, and that initial stream placing exerts a potent influence on the child's development. Plowden (1967) says that it is desirable to assess the child before he leaves primary school, but also states that where selection is made, it should cease to be based on an " externally imposed battery of intelligence and attainment tests."The evidence of all the research described and referred to here is that schools generally are, as yet, unable to make accurate assessments of ability without the aid of such " externally imposed batteries of tests." This lends weight to Plowden's further comments on the need for the development of test and assessment procedures for use within the school.

VI.—REFERENCES.

ARMSTRONG, H. G. (1966). A comparison of the performance of summer- and autumn-born children at 11 and 16. *Brit. J. Educ. Psychol.*, 36, 72-76.

FREYMAN, R. (1965). Further evidence on the effect of date of birth on subsequent school performance. *Educ. Res.*, 8, 58-64.

JACKSON, B. (1964). *Streaming: An Educational System in Miniature.* London: Routledge and Kegan Paul.

JINKS, P. C. (1964). Investigation into effects of date of birth on subsequent school performance. *Educ. Res.*, 6, 220-225.

LUNN, J. C. B. (1967). The effects of streaming and other forms of grouping in junior schools. *New Res. in Educ.*, 1, 4-75.

PLOWDEN (1967). *Children and their Primary Schools*: A Report of the Central Advisory Council for Education (England). London: H.M.S.O.

(*Manuscript received* 10*th August*, 1966)

ROAD SAFETY EDUCATION FOR YOUNG CHILDREN

By KENNETH PEASE AND BARBARA PRESTON
(*Child Study Unit, Dept. of Psychology, University of Manchester*)

SUMMARY. Trends in road casualty figures among children in recent years are evaluated and interpretations suggested. The experiment reported here shows that the educator's conception of the function of the kerb drill, i.e., to detect traffic, is not reflected in its function as perceived by many young children, who in the experimental setting behave as though they regarded the performance of the kerb drill as in itself sufficient to ward off the dangers of the road. The experiment also demonstrates that the notions of directionality which the kerb drill presupposes are not those of 5-6-year-old children, who often regard ' left ' and ' right ' as parts of the environment rather than relations to the person. The ' static ' conception of directionality becomes more frequent after exposure to a propaganda film in which the kerb drill exemplars perform facing their audience.

I.—INTRODUCTION.

" When you cross the road by day or night,
Beware of the dangers that loom in sight.
Look to the left, and look to the right,
Then you'll never, never get run over."*

LAST year 900 children under the age of 15 were killed on the roads of Great Britain, 13,962 were injured seriously enough to be detained in hospital and 47,139 were slightly injured.[9] The proportion of child deaths resulting from road accidents is all the more striking after the virtual elimination of infectious diseases which formerly caused many such deaths. In England and Wales in 1964, 21 per cent. of deaths of children aged between 5 and 10 years were due to neoplasms of one kind or another, 20 per cent. to road accidents and only 4 per cent. to infectious diseases.[6]

Of the three E's of Road Safety, Education, Engineering and Enforcement, the emphasis in this country has always been on education, whereas, in the U.S.A., engineering and enforcement have been more strongly advocated. All three E's are important, but for children road safety education is vital. What evidence have we, however, of the effect of road safety education ? In 1930 there were 1,685 deaths of children under 15 on the roads.[8] The comparable figure in 1965 was 900. Six reasons for this drop other than the reason of safety education are suggested.

(1) Advances in medical science in the past thirty-five years mean that children who would have died as a result of their injury now survive.

(2) School dinners cut the risk. Many children now only cross busy roads twice a day, whereas previously they would have had to cross them four times daily.

(3) A road death was redefined in 1954. Before then a child who died as a result of a road accident was classified as a road death, but from 1954 a seriously injured child who lived for more than thirty days was not so classified. The Road Research Laboratory[7] considers that about 5 per cent. of people killed ' on the roads ' survive for more than thirty days. Child deaths dropped 17 per

---

* By kind permission of Campbell Connelly and Co., Music Publishers.

cent. the year the new classification was introduced, so it seems possible that the effect of the new legislation could have been particularly large for children.

(4) Child road deaths decreased by 9·5 per cent. after the 1930 Road Traffic Act. This Act stipulated minimum age limits for drivers, introduced the offences of 'careless driving' and 'driving under the influence, etc.', made third party insurance cover obligatory and introduced the penalty of disqualification.

(5) Child road deaths decreased again by 13·5 per cent. after the 1934 Road Traffic Act. This Act introduced 30 m.p.h. speed limits in built-up areas, driving tests and pedestrian crossings with Belisha Beacons.

(6) Child pedestrian deaths decreased by 13·5 per cent. after the introduction of Zebra crossings in 1952.

Reducing 1,685 by these successive percentage reductions would bring the present figure to 950, but this is an obvious oversimplification. There are thus many reasons other than road safety education for the drop in child deaths on the roads since 1930.

Looking more closely at the picture since 1945, the trends in children killed and children injured are quite different from each other,[1,3,6] (see Figures 1 and 2). The number of children injured has risen consistently, as would be expected given the increased use of motor vehicles. The estimated motor fuel consumption per year is plotted on figure 2 to allow comparison with number of children injured. As for child deaths on the road, the downward trend evident before 1960 has now been reversed. The writers very tentatively relate this change at this time to the speeding up of urban traffic (92 per cent. of child pedestrian injuries occur in built-up areas) following the Road Traffic Act, 1956, and the Road Traffic and Roads Improvement Act, 1960. Eaton[3] showed that speeds in central London, after decreasing from 1947 to 1958, began to increase again from 1960.

Thus, in view of changes in road casualty figures among children induced by legislative and other modifications, changes attributable to road safety education are not apparent.

One legitimate method of investigation of road safety education, and one which assumes greater importance in the absence of adequate evidence from road casualty figures, is to study the process at source, i.e., to determine the extent to which the messages of road safety are communicated to the child by the educator, and the interpretation which the child places upon the information to which he is exposed.

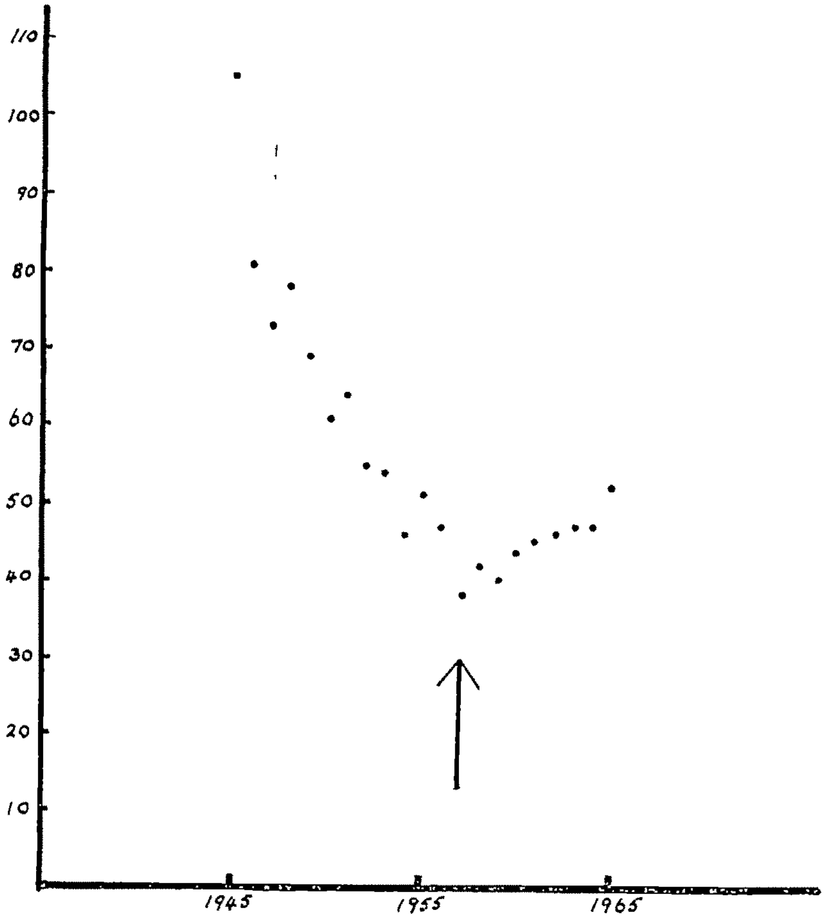
A major tool in the armoury of the road safety educator is, and has been for the last twenty-five years, the 'kerb drill,' a prescribed series of actions faithful adherence to which will ensure, for the child or adult who wishes to cross the road, detection of approaching vehicles and consequent choice of a safe time to cross. The child is encouraged to perform the drill, with the tacit assumption that he regards it as the most efficient means of recognising hazards and acting appropriately to their presence or absence. The suspicion that the child may see the drill as something other than a mere instrument of detection may be aroused by consideration of the world view of the young child as described by Piaget[4,5] in which rules are given the status of physical laws. The experiment reported here sets out to investigate the relationship between the kerb drill and safety as conceived by the young child, the hypothesis being that the child will consider mere production of the drill to ensure safety.

A second matter of concern in this study was the effect of the directional ideas of the young child upon his educability with regard to the kerb drill.

It is well known that the young child often considers left and right to refer to parts of the environment rather than as relations to himself. The effect of this 'static' conception of direction upon comprehension of the kerb drill is a matter of interest in itself. In addition, the road safety propaganda films which the writers have seen generally portray the kerb drill exemplars as facing their audience. This experiment sets out to determine the influence of such a propaganda film upon the directional ideas of young children as revealed in their use of the kerb drill.

FIGURE 1

CHILD PEDESTRIANS KILLED PER MILLION CHILDREN EACH YEAR SINCE 1945

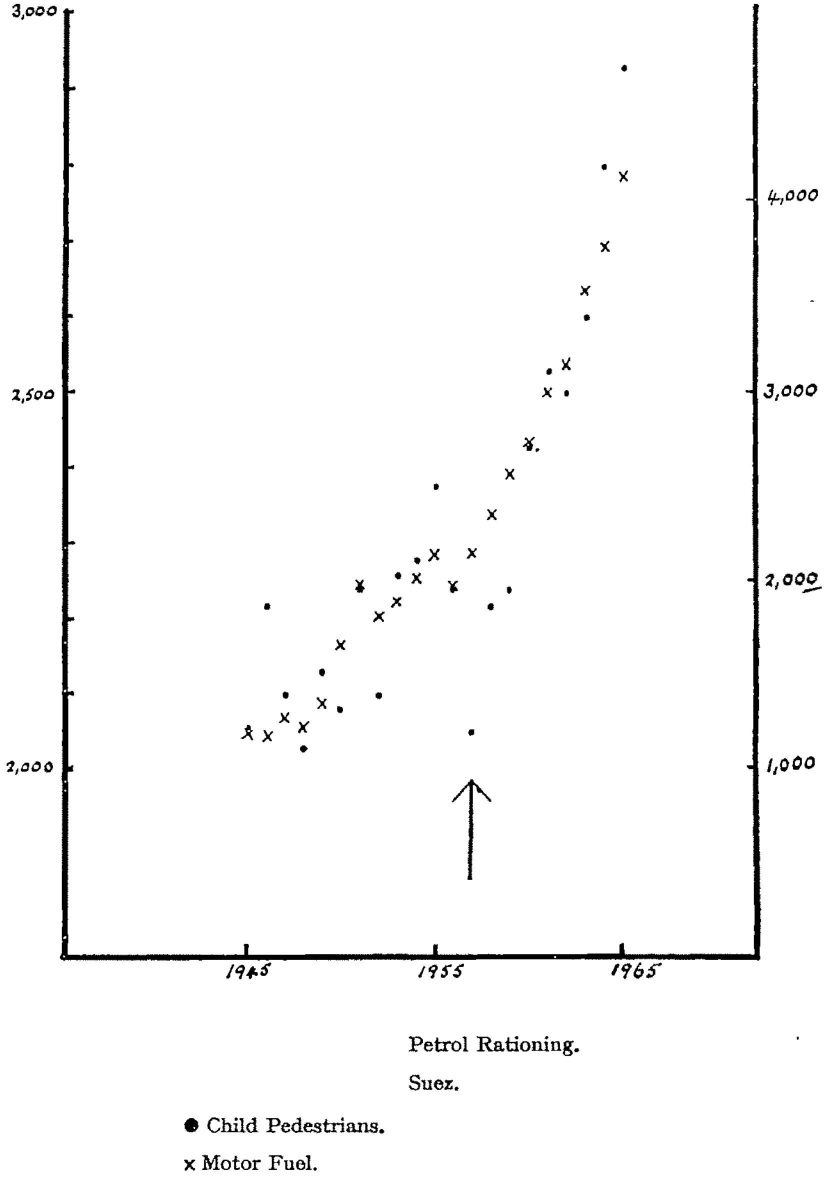


Petrol Rationing.

Suez.

FIGURE 2

CHILD PEDESTRIANS INJURED PER MILLION CHILDREN AND MOTOR FUEL CONSUMED EACH YEAR SINCE 1945.



II.—PLAN OF THE INVESTIGATION.

(i) *Subjects.*

Subjects were 70 children from two state infant schools in Manchester, situated close to each other and with a very similar intake in terms of socio-economic status and ethnic composition. The experimental group comprised 42 children, 21 boys and 21 girls, mean age 5 yrs. 6 mths., S.D. 5.2 mths., from one of the schools. The control group, from the other school, numbered 28, 14 boys and 14 girls, mean age 5 yrs. 8 mths., S.D. 4.6 mths. The two groups were divided as to school in this way because of the difficulties which would have been involved in showing the propaganda film to some of the children in a school and not others. It is felt that the groups are comparable in all relevant respects.

(ii) *Experimental Procedure.*

In the experimental situation, children were seen individually in whatever room was convenient for the school and suitable for the experimenter. Children were put at their ease as far as was possible and were then asked to pretend that a defined space of the floor (which was clearly marked) was a road over which they had to cross safely. They were asked to say and do what they ought to before crossing. In most cases this was enough to elicit the kerb drill. Where it was not, broader hints were given until eventually some variant of the procedure at the kerb was produced. (Two Indian girls were unable to understand English and are not considered in the results.) When the kerb drill was completed, each child crossed the 'road' and was asked to turn round to face the experimenter, and was then asked to do and say the kerb drill again so that he could cross the road back to the experimenter.

If the direction described as 'right' by the child differed according to his position in relation to the experimenter, then 'right' was considered to denote part of the environment rather than a relation to the child. If he looked to his right on both occasions, then his notion of direction was considered to be a relational one.

At this stage in the experimental procedure, the child was once again asked to do and say the kerb drill to cross the 'road' away from the experimenter. After he had looked in one direction, the experimenter asked him to suppose that there was a car coming, and asked what he would do in this situation. The child's responses were classified according to whether he (*a*) waited for the car to go and continued his kerb drill from where he had been interrupted, (*b*) simply crossed the road after the car had gone without mention or sign of further attempts to detect more traffic, or (*c*) made sure that the road was clear before crossing, whether or not this took the form of the conventional kerb drill. If response (*a*) or (*b*) resulted, the child was given counter-suggestions. Those who when pressed gave a response type (*c*) were classified as giving an adequate response 'under pressure.' The rationale for this classification was that responses type (*a*) or (*b*) which were not changed under pressure indicated that the child did not comprehend the real function of the kerb drill, since when 'detection' of a car occurred, subsequent behaviour would not have guaranteed safe passage across the road. Response types (*a*) and (*b*) are profoundly different from each other but are combined in the results section because they are both inadequate to the purpose of the child's safety.

Both experimental and control groups underwent the experimental procedure twice, with an interval of four days. On the day of the experimental group's second session and before it, the experimental group was shown the road safety propaganda film "Mary had a Little Lamb," produced by the Petroleum

Films Bureau, and which has been popular in use with children of 5–7 years for over a decade. This is in many ways a delightful film, telling the story of a school girl's pet lamb who nearly gets run over and then goes and learns the kerb drill at school, and is afterwards safe while crossing the road. In common with other films of this kind, however, the kerb drill exemplars in ' Mary ' face their audience as they look left and right, and hence children without well-established relational notions of direction may be induced,under such circumstances, to readopt a static conception of direction, since the verbalisation of direction in the film does not coincide with the direction which a relational child would attribute to the exemplar's action.

After this film had been shown to the experimental group, the second experimental session was completed. The second session of the control group occurred without the intervening film. Both schools assured the experimenter (who was the same person for both groups) that no relevant instruction was given between the experimental sessions.

The artificiality of the experimental situation must not be lost sight of regarding the findings here reported, particularly considering the advent of the supposed car during the kerb drill. However, children were not apparently disturbed by the interruption, and for the most part readily said what they would do next. The writers are satisfied with the validity of the results obtained an believe that the findings do have meaning for road safety education. They would be more than hestitant, however, to draw too close a comparison between a child's behaviour in the experimental situation and his behaviour on the road.

III.—Results.

TABLE 1

The Experimental Group's Kerb Behaviour When a ' Car ' Has Passed.

	Pre-film Trial 1	Post-film Trial 2	
Adequate precaution taken when car gone	7	9	16
Adequate precaution under pressure when car gone	17	18	35
No adequate precaution when car gone	18	15	33
Total	42	42	84

Table 1 shows that on a second trial, after being shown the road safety propaganda film " Mary had a Little Lamb," the experimental group showed only a marginal improvement in precautions taken before crossing the ' road ' after an imaginary car has interrupted the performance of the kerb drill.

TABLE 2

The Control Group's Behaviour When a ' Car ' Has Passed.

	Trial 1	Trial 2	
Adequate precaution taken when car gone	4	6	10
Adequate precaution under pressure when car gone	10	12	22
No adequate precaution when car gone	12	8	20
Total	26	26	52

Table 2 shows that on a second trial, without being shown the propaganda film between trials, the control group also showed a marginal improvement in precautions taken before crossing the ' road ' after an imaginary car has interrupted the performance of the kerb drill.

TABLE 3

EXPERIMENTAL GROUP'S DIRECTIONALITY BEFORE AND AFTER PROGAGANDA FILM.

		Trial 2. Post-film		
		Relational	Static	
Trial 1 Pre-film	Relational	5	12	17
	Static	4	18	22
	Total	9	30	39

Table 3 shows children's conception of direction before and after the prop ganda film.

Three subjects were not considered here because of inconsistency within an experimental session.

Using the McNemar test for the significance of changes, $\chi^2=3\cdot06$ $p<\cdot10$. Thus, there is a tendency for the experimental group to change from a relational to a static conception of direction after being shown the road safety propaganda film " Mary had a Little Lamb," in which kerb drill exemplars face their audience.

TABLE 4

CONTROL GROUP'S DIRECTIONALITY ON FIRST AND SECOND TRIALS.

		Trial 2		
		Relational	Static	
Trial 1	Relational	4	2	6
	Static	9	10	19
	Total	13	12	25

Table 4 shows the change in conception of direction of the control group between trials.

One subject was not considered for the same reason as the three omissions in the experimental group. Using the McNemar test for the significance of changes, $\chi^2 3\cdot27$ $p<\cdot10$. Thus, there is a tendency for the control group to change from a static to a relational conception of direction on a second trial.

In Table 5 we compare changes in the two groups between trials directly. Only children who changed their conception of direction between trials are entered here. Using the Fisher-Yates Exact Test, $p=\cdot005$. Thus, the two groups differ with respect to the direction of change of notions of direction. Where change occurs between the two trials in the two groups, the direction

TABLE 5

COMPARISON OF CHANGES IN CONCEPTION OF DIRECTIONALITY IN THE TWO GROUPS.

	Relational-Static Changes	Static-Relational Changes	
Experimental Group	12	4	16
Control Group	2	9	11
Total	14	13	27

of that change is different in the two groups, and the incidence of the two directions of change in the two groups statistically distinguishes the groups.

Verbalisation.

Defined as likelihood of detection of approaching vehicles, the adequacy the child's verbalisation of the kerb drill was assessed. Despite the vagueness of this, agreement between two independent judges of this adequacy occurred in thirty-eight cases out of forty-two. It appears that the criteria of adequacy were (*a*) stating " Stop at the kerb," (*b*) verbalising right immediately before terminating the drill, (*c*) alternation between directions. In cases where all these were equivalent, sheer length of the drill was taken into account.

Using the classification produced on the performance of children in the experimental group, eighteen of the forty-two children produced a better verbalisation of the drill before the film, twenty-four gave a better verbalisation after it, and in four instances the verbalisation was equally good before and after. Dividing the ' equal ' category between the other two groups, we have a split of 26–20. Applying a sign test (one tail) this is not statistically significant.

In the control group of twenty-six children (inter-judge agreement in twenty-three cases), seven children did better on the first trial, nine did better on the second trial, and ten performed identically on both occasions. By inspection, this split does not approach statistical significance.

IV.—DISCUSSION.

From the evidence here presented, it appears that young children often do not behave as though they regarded the kerb drill to be a device for detecting danger. Most of the children, when interrupted while performing their kerb drill by a supposed car, do not spontaneously resume after the ' car' has passed in such a way that detection of other vehicles would occur, and a large number even when encouraged to expand their performance, do not do so. The interpretation suggested here is that the child feels that merely undertaking the kerb drill obviates the need for further appraisal of the situation when cars have passed. It is felt that the way in which the child undertakes the drill in the experimental situation lends support to the interpretation that the kerb drill is seen as merely a safety rite to be gone through before crossing the road, in that the child's movements appeared anything but natural, and the eyes were directed right and left without apparently focussing on what was there. The way in which many of the children walked across the ' road ' during the experiment was for the most part extraordinarily stylised. Those who have watched the kerb drill behaviour of young children on the public roads may agree that this observation is not due to an artefact of the experimental situation.

As for the directionality aspect of the study, it is to be noted that after seeing the film " Mary had a Little Lamb," in which Mary and her lamb perform their kerb drill facing their audience, there was a tendency for the experimental group to become more static in its directionality. This is more striking in contrast with the opposite tendency of the control group. The writers are assured that this was not the result of teaching between the sessions. The extent of the difference in direction of change between the two groups is statistically highly significant.

Ability to verbalise the kerb drill did not improve to a statistically significant degree in either group.

V.—CONCLUSIONS.

The kerb drill is not always used by young children as though they regarded it merely as an efficient detector of approaching traffic. Often interruption of the kerb drill by a supposed car will produce a resumption of behaviour which would not ensure safety. This phenomenon was not influenced by the showing of the road safety propaganda film " Mary had a Little Lamb."

To many young children, ' right ' and ' left ' refer to parts of the environment rather than relations to the self. This necessitates care in the use of the kerb drill by road safety educators. The ' static ' notion of direction tends to become more frequent among children who have been shown " Mary had a Little Lamb " in which kerb drill exemplars face the audience, whereas a control group tends more frequently to evidence a relational idea of direction on a second trial.

The showing of " Mary had a Little Lamb " did not significantly affect the ability of children to verbalise the kerb drill, which ability improved marginally from the first to the second trial in both experimental and control groups.

ACKNOWLEDGEMENTS.—The writers wish to thank Messrs. McFarlane and Cheetham of Manchester Road Safety Offices, and the head teachers and teachers who co-operated in the research reported.

VI.—REFERENCES.

[1] BRITISH ROAD FEDERATION (1966). *Basic Road Statistics.*

[2] EATON, J. E. (1964). *London Traffic Survey.* Road Research Laboratory Report No. 5. Harmondsworth : Road Research Laboratory, 1966.

[3] *Monthly Digest of Statistics* (1946-66). H.M.S.O.

[4] PIAGET, J. (1929). *The Child's Conception of the World.* New York : Harcourt, Brace.

[5] PIAGET, J. (1932). *The Moral Judgment of the Child.* London : Kegan Paul.

[6] *Registrar-General's Statistical Review for England and Wales* (1964). Part I, Medical. H.M.S.O.

[7] ROAD RESEARCH LABORATORY (1963). *Research on Road Safety.* H.M.S.O.

[8] ROYAL SOCIETY FOR THE PREVENTION OF ACCIDENTS (1957-1964). *Road Accident Statistics.* Terminal House, 52, Grosvenor Gardens, S W.1.

[9] ROYAL SOCIETY FOR THE PREVENTION OF ACCIDENTS (1966). *Road Accident Statistical Review* (March). Terminal House, 52, Grosvenor Gardens, S.W.1.

(*Manuscript received* 18*th August*, 1966)

EFFECTS OF OMITTING BRANCHES AND QUESTIONS FROM A SCRAMBLED TEXT

By K. D. DUNCAN* and T. GILBERT†
(*Army Personnel Research Establishment, West Byfleet*)

SUMMARY. Three versions of a programmed book (*The Waterloo Campaign*) were studied by groups of 15-16-year-old boys. The versions were: the original *branching* programme; *prime* path frames only (branches omitted); continuous *text* (multiple choice questions and branches omitted). The continuous *text* was associated with slightly superior immediate post-test scores and slightly inferior delayed post-test scores, compared with the other two versions. These differences were not statistically significant; nor were differences in attitude to the three versions. Study time was less in the *prime* than in the *branching* version and less than either of these in the *text* version. Study time differences were substantial and statistically significant. With all three versions the correlation between intelligence and attainment was statistically significant: it was highest in the *text* version and lowest in the *branching* version.

I.—INTRODUCTION.

WIDELY differing styles of programmed instruction have been advocated[2, 6, 7, 11] One approach to their empirical evaluation is to take an existing programme, which may be presumed to be a fair realization of a particular style, and deliberately to remove those features which are alleged to be critical by proproponents of the style in question. This has been done in the case of a Skinnerian programme by replacing missing words and condensing the content,[8] in the case of a Crowderian programme by removal of 'branches'[5, 10] and in the case of a 'mathetics' programme by omitting 'prompt' and 'release' frames.[13]

In general these changes seem not to have adversely affected attainment, although only one of the foregoing studies included delayed testing to assess retention.[13] A possible interpretation of these findings is that the major contribution to learning lies more in general programming procedures, such as deriving and specifying terminal behaviour, structuring the material to be mastered, etc., than in specific features of programming style.

II.—THE PRESENT EXPERIMENT.

(i) *Contribution of 'Branches' and Multiple Choice Questions.*

The present experiment was stimulated by a report of work by Senter and his associates, in which they deleted the remedial branches from Crowder's scrambled text, 'The Arithmetic of Computers,'[1] without apparently affecting attainment.[10] The authors warn against generalising this rather surprising finding, pointing out that the control group, who used the unmodified text, may not have made much use of the branches. They suggest that 'younger or less qualified' subjects than their own (17–22-year-old college students), would possibly have benefited more from the remedial sequences. Essentially the same finding has been reported by Kaufman,[5] who also used 'The Arithmetic of Computers.'

---

* Now at Department of Psychology, University of Hull.
† Now at National Foundation for Educational Research.

In both experiments, the presence of branches, apart from having no effect on attainment, had no significant effect on study time. This latter finding is in line with the interpretation that those subjects who had access to remedial sequences did not expend much effort on them. Neither experiment carried out delayed post-tests to assess retention.

Accordingly, in the experiment reported here, both immediate and delayed post-tests were used to assess the contribution of branches to attainment. Also subjects were chosen who were not so familiar with the material, or so intelligent, that they might be expected to do well despite any variations in presentation. Finally, in addition to deleting branches, the more drastic modification of removing the multiple choice questions to make a continuous text was incorporated in the experiment.

From this last modification can be inferred the contribution to attainment of any 'covert responses' to the multiple choice questions. The particular comparison of this learning condition with simply reading continuous text is one which has apparently been neglected, most work having been with *constructed* rather than *multiple choice* responses and mainly devoted to the issue of *overt versus covert* responding.[4] An experiment by Williams, using modifications of a linear programme, did include a reading condition and a two-choice responding condition, but the latter (which produced superior attainment) required *overt* responses.[14] In another experiment, multiple-choice branching programmes were compared with continuous text, but in this case effects of responding cannot be separated from effects of the branches.[3] One interesting finding of this last investigation was an interaction with subject matter—the text being more efficient for instruction in 'military rules and codes of conduct,' but not in the case of First Aid.

(ii) *Material.*

The programme used, 'The Waterloo Campaign,'[12] is a novel and, in the writers' view, entertaining application of the multiple choice branching technique. After each frame the reader has to decide what would be the next action of a particular leader at that stage of the campaign. Branches serve the usual role of repetition and further explanation.

In addition to the original, subsequently referred to as 'branching version,' two other versions were prepared:

(*a*) A version from which the branches were removed, subsequently referred to as 'prime version,' and in which each prime path frame, with its multiple choice question, was on a separate sheet so that the subject had to turn over to see the next one.

(*b*) An uninterrupted text, in which the original branches and questions were removed and the prime path frames were fitted into a continuous narrative, with as slight textual modification as was needed to make sense of the transitions ('text version').

The two modified versions were produced in the same print and format as the original.

(iii) *Procedure.*

The subjects were 15–16-year-old boys of the Junior Leaders Regiment, Royal Artillery. On each of four successive mornings a class of fifteen–twenty boys was made available for testing and study of the programme. These sessions were supervised by one of the authors.

The subjects were given test AH4 followed by a short, written answer pre-test to assess their existing knowledge of the subject matter. After a fifteen-minute break, they were assigned one of the three versions of the programme by randomising the class list. They were told they would be tested afterwards and were urged to read carefully and try to remember personalities, decisions and events described in the text. Each subject informed the investigator as soon as he had finished the programme, at which time he was given the post-test.

Data were initially obtained for seventy-two subjects. However, when the post-test was subsequently re-administered after approximately twenty weeks, fourteen subjects were not available for testing (e.g., because they were ill, or had left the unit, etc.). Data for these fourteen subjects were thus incomplete and were, therefore, discarded.

The post-test was untimed and consisted of thirty-five questions, three in multiple choice and the remaining thirty-two in short answer form. It included all twelve pre-test questions. The scoring key distinguished information which could be obtained : (*a*) only from prime path frames (P score) ; (*b*) *either* from prime path frames *or* branches (P/B score). The test was written and all test papers marked by one of the authors. Since the marking of most of the questions was not objective, it was carried out ' blind,' i.e., the names and other particulars of the subjects which might have identified which version they had studied were removed from the test papers before marking.

The post-test ended with eight multiple choice questions on the subjects' opinions about the programme's clarity, their ease or difficulty in understanding it and the method of presentation. The five alternative answers to each question were assigned values of 1 to 5, according to their favourability. The sum of the value of each subject's choices was taken as a ' positive attitude score.'

III.—Results.

Means and standard errors for the three groups are presented in Table 1. In terms of the associated standard errors, the differences between groups in attainment (the six post-test criteria) and attitude score are statistically insignificant, whereas the differences in study time are substantial and highly significant statistically.

With regard to matching, the differences between the groups in AH4 and pre-test scores do not reach the 10 per cent. level of significance. Nevertheless, the matching is not perfect, and accordingly, a linear regression analysis was carried out in the case of each of the dependent variables, i.e., attainment, attitude score and study time, using AH4 and pre-test scores as independent variables. (The procedures followed are described by Quenouille, 1952, Chapter 5; his Table 5·4f gives the sources of variation identified.) In each case, a statistically significant regression was found on either AH4, or on the pre-test, or a significant multiple regression on both. These analyses confirmed the insignificance of the differences between the groups in attitude score and the six attainment variables, as well as the significance of the study time differences ('adjustments ' to the means in this last case were negligible).

Table 2 gives the intercorrelations between intelligence, study time and attainment as measured by the immediate and delayed post-test total scores. The total score test re-test coefficients indicate that the scoring was fairly reliable. First order partial coefficients for the intercorrelations of AH4, study time and immediate post-test are also shown, i.e., the correlation between each pair is shown with effects of the third variable partialled out.

TABLE 1

MEANS, WITH ASSOCIATED STANDARD ERRORS IN BRACKETS, FOR INTELLIGENCE, PRE- AND POST-TESTS, STUDY TIME AND ATTITUDE SCORE.

		Group	
	Branching n=18	Prime n=21	Text n=19
AH4	68·0	71·7	74·7
	(4·16)	(3·02)	(2·87)
Pre-Test	1·3	2·0	2·1
	(0·32)	(0·46)	(0·47)
Study Time (minutes)	54·2	46·0	33·7
	(3·01)	(2·12)	(1·79)
Attitude Score	13·6	12·6	14·5
	(1·09)	(0·73)	(0·95)
Immediate Post-Test:			
P score	2·9	3·1	4·1
	(0·45)	(0·52)	(0·57)
P/B score	8·3	8·3	8·4
	(1·08)	(0·72)	(1·10)
Total score	11·2	11·4	12·5
	(1·38)	(1·10)	(1·41)
Delayed Post Test:			
P score	0·9	0·6	1·0
	(0·21)	(0·16)	(0·27)
P/B score	5·0	5·6	4·5
	(1·01)	(0·94)	(0·76)
Total score	5·9	6·2	5·5
	(1·15)	(0·95)	(0·97)

P score : Information which could only be obtained from prime path frames.
P/B score : Information which could be obtained *either* from prime path frames *or* branches.

IV.—DISCUSSION.

Although the programme used in this experiment was short, the branches and the multiple choice questions clearly *did* entail marked differences in study time on the part of the subjects. This finding contrasts with the failure to find significant time differences in two previous investigations[10, 5].

Nevertheless, the extra study time involved by the branches and the multiple choice questions appears not to have benefited attainment. Furthermore, retention over a twenty-week period which, more than an immediate test, might be expected to show up the benefits of these programming variables, failed to do so. If, indeed, branching and multiple choice questions assist learning, this should be most evident in the P/B scores: there is a slight tendency, in this direction, in the case of the delayed post-test means, but it is not statistically significant. The foregoing indications then, taken with the absence of significant differences between the three versions in attitude scores, together indicate that the text version is the most efficient.

With regard to the intercorrelations observed, the significant relation between study time and intelligence, found with the branching but not the other versions, is what one might expect. Selecting the incorrect alternative answer and reading a branch as a result, would increase study time ; the frequency of such excursions to branches must be partly determined by intelligence.

TABLE 2

INTERCORRELATIONS OF INTELLIGENCE, STUDY TIME AND POST-TEST SCORES.

	AH4	Study Time	Immediate Post-Test
Branching Version, n=18			
Study Time	—0·73** **—0·76****		
Immediate Post Test ..	0·33 **0·43**	—0·06 **0·28**	
Delayed Post Test	0·37	—0·01	0·78**
Prime Version, n=21 :			
Study Time	—0·18 **—0·16**		
Immediate Post Test ..	0·50* **0·50***	—0·07 **0·02**	
Delayed Post-Test	0·36	—0·14	0·65**
Text Version, n=19 :			
Study Time	—0·36 **—0·15**		
Immediate Post-Test ..	0·67** **0·62****	—0·39 **—0·22**	
Delayed Post-Test	0·48*	—0·35	0·80**

The bold figures are first order, partial coefficients for the inter correlations of AH 4, Study Time and Immediate Post-Test.
* p<0·05. ** p<0·01

It is noteworthy that the correlation between the immediate post-test and intelligence of 0·62 in the case of the text, reduces to 0·50 in the presence of multiple choice questions and further reduces to 0·43 in the version with branches. This seems a fairly clear indication that the introduction of programming variables diminishes the extent to which attainment depends on intelligence—unfortunately, in this case, at the expense of overall increases in mean study time.

V.—CONCLUSION.

The findings of this experiment further support the view that specific features of programming style are relatively unimportant to learning, once the material to be learned has been adequately ordered and structured.

Specifically, the presence of multiple choice questions and branches requires extra study time and is not associated with improved learning, or with significantly improved retention after a delay of twenty weeks. However, the presence of these programming features does, apparently, diminish the extent to which attainment is dependent on intelligence.

VI.—REFERENCES.

[1] CROWDER, N. A. (1960). *The Arithmetic of Computers*. New York : Doubleday. London : E.U.P.

[2] CROWDER, N. A. (1963). Intrinsic programming : facts, fallacies and future. *Prospectives in Programming*. (Edited by R. T. Filep.) London : Collier-Macmillan.

[3] FOLLETTIE, J. F. (1961). *Effects of Training Response Mode, Test form and Measure of Acquisition of Semi-ordered Factual Materials*. Research Memorandum. Alexandria, Va. : Human Resources Research Office.

[4] HOLLAND, J. G. (1965). Research on programming variables. *Teaching Machines and Programmed Learning* II—*Data and Directions.*, (Edited by R. Glaser). Washington: National Education Association.

[5] KAUFMAN, R. A. (1963). An experimental evaluation of the role of remedial feedback on an intrinsic program. *Journal of Programmed Instruction*, 2, 4, 21-30.

[6] PENNINGTON, D. F., and SLACK, C. W. (1962). The mathetical design of effective lessons. *Applied Programmed Instruction.* (Edited by S. Margulies and L. D. Eigen.) New York: Wiley.

[7] PRESSEY, S. L. (1963). Teaching machine (and learning theory) crisis. *J. Appl. Psychol.*, 47, 1-6.

[8] PRESSEY, S. L., and KINZER, J. R. (1964). Auto-elucidation without programming! *Psychology in the Schools.* 1, 4, 359-365. Reprinted. Paris: Editions Gamma.

[9] QUENOUILLE, M. H. (1952). *Associated Measurements.* London: Butterworth.

[10] SENTER, R. J., NEIBERG, A., ABAMA, J. S., and MORGAN, R. L. (1964). An evalaution of branching and motivational phrases in a scrambled book. *Programmed Learning.* 1, 124-133.

[11] SKINNER, B. F., and HOLLAND, J. G. (1960). The use of teaching machines in college instruction. *Teaching Machines and Programmed Learning.* (Edited by A. A. Lumsdaine and R. Glaser) Washington. National Education Association.

[12] THORNHILL, P (1965). *The Waterloo Campaign.* London: Methuen.

[13] WALLIS, D., DUNCAN, K. D., and KNIGHT, M. A. G. (1966). *Programmed Instruction in the British Armed Forces*, pp. 43-44. London: H.M. Stationery Office.

[14] WILLIAMS, J. P. (1963). A comparison of several response modes in a review program. *J. Educ. Psychol.*, 54, 253-60.

(*Manuscript received 2nd August*, 1966)

THE CAUSES OF FAILURES IN CHILDREN TO SORT BY TWO DIFFERENT DIMENSIONS ON SUCCESSIVE TRIALS

By P. E. BRYANT

(Institute of Psychiatry, Medical Research Council, Maudsley Hospital, London, S.E.5)

Summary. Two hypotheses were examined as explanations of failures in children of 5 and 6 years of age, to change from sorting cards in one of two dimensions, either colour or shape, to sorting in the other dimension. The first hypothesis was that these failures are the result of a learned set to respond on the basis of one dimension, and the second was that they are caused by the subject's learning not to respond on the basis of particular stimulus aspects. The results of an experiment with two different types of material gave support for the second hypothesis.

I.—Introduction.

Studies by Thompson (1941), Reichard, Schneider and Rapaport (1944) and Inhelder and Piaget (1964) have shown with children between 5 and 10 years that the ability to sort the same material in two successive trials by two different dimensions increases with age. These studies utilised adaptations of the colour-form sorting test developed by Goldstein and Scheerer (1944), in which material characterised by different colours and shapes is used. In this test the subject is required to sort the material by one of the two possible dimensions, e g., colour, and then to sort the same material by the other dimension, e.g., shape. Changing to the second dimension is described as shifting.

All these studies used what can be described as non-hierarchical material, in that each colour was represented in all the shapes, and each shape in all the colours. With this material sorting by, for example, colour involved separating the colours and placing the shapes together. Thus, if the colours were red and blue and the shapes, circles and crosses, sorting these cards on the basis of their colours would involve making a pile of red cards and a pile of blue cards, each pile consisting of crosses and circles. The fact that such material was used in these studies leads to the possibility of two different hypotheses to explain failures by younger children to shift to the second dimension.

The first hypothesis is that failure to shift is the result of a general set, which is built up during the first sorting, to concentrate on one dimension and to ignore the others. According to this hypothesis, sorting on the basis of colour creates in younger children a strong set to concentrate on colour and to ignore other dimensions, and as a result, the subject fails to shift to shape.

The second hypothesis is that failure to shift is caused by the nature of the material, rather than by set. According to the second hypothesis, the subjects learn during the first sorting that specific stimulus aspects are irrelevant, and this learning prevents them utilising the second dimension. In the above example, sorting cards into red and blue involved making separate red and blue piles and placing circles and crosses together in the same piles. The second hypothesis is that the experience of mixing circles and crosses together leads to learning that these stimulus aspects are irrelevant, and that as a result, the subject is unable to utilise circles and crosses as cues for a later sorting on the basis of shape.

The distinction between these two hypotheses is important if one is to explain the nature of the cognitive processess which lead to failures to shift. The first hypothesis would lead to the prediction that younger children would fail to shift with every kind of material that is characterised by more than one dimension. The second hypothesis would predict such failures only with non-hierarchical material of the kind described above. On the other hand, if material is used in which sorting on the basis of, for example, colour does not involve placing the different shapes together in the colour piles, the second hypothesis would not predict failures to shift. Such material can be described as hierarchical material. If, for example, the material consisted of two shapes, circles and crosses, and four colours, red, blue, green and yellow, it would be hierarchical if the circles were represented only in the red and blue, and the crosses only in green and yellow. In this case, sorting on the basis of colour would involve sorting into colour piles of red, blue, green and yellow, and in none of these piles would the two shapes, circles and crosses, be mixed together.

With this material the first hypothesis would predict failures to shift to shape after sorting by colour, because the subject would have built up a general set to concentrate on colour. However, the second hypothesis would not predict a failure to shift to shape after sorting by colour because colour sorting would not have involved learning to ignore the shapes by mixing them together in the different colour piles.

In order to examine the relative validity of these two hypotheses as explanations of failures to shift in younger children, an experiment was designed in which 5 and 6-year-old children were made to sort material of different colours and shapes, first by one dimension and then by the others. Some of the children were given hierarchical material, others non-hierarchical material. It was claimed that if the children made an equal number of failures to shift with both kinds of material, this would be evidence that failure to shift were the result of a general set being built up during the first sorting. If, on the other hand, there were more failures to shift with non-hierarchical than with hierarchical material, this would be evidence that failures to shift are caused by previous learning to avoid specific stimulus aspects.

II.—METHOD.

The subjects consisted of ninety-six English-speaking children drawn from two primary schools in Geneva. These were divided into two groups of forty-eight subjects, the groups differing in terms of their mean age levels. The mean age for Group I was 67 months, range 59-72 months and for Group II 80 months range 74-87 months. These groups were subdivided into three subgroups of sixteen subjects each. The mean age for these three subgroups, A, B and C, was 66 months, 67 months and 67 months, respectively, for Group I, and 81 months, 79 months and 80 months, respectively, for Group II.

The material used in the experiment is described in Table 1.

It can be seen from Table 1 that there were three sets of material and that each set consisted of four different cards, each card being reproduced ten times. The first set (NH) was the non-hierarchical set in which the four cards were formed from two colours and two shapes, each colour being represented in both the shapes, and each shape being represented in both the colours. The second set (H1) consisted of hierarchical material in which there were two shapes and four colours. This was described as material in which colour was the major dimension and shape the minor dimension. Each shape was represented in two of the colours, and each colour in only one shape. The third set (H2) of material was

TABLE 1

THE THREE SETS OF CARDS.

(NH) Non-Hierarchical	
10 blue squares 10 red squares	10 blue circles 10 red circles
(H1) Hierarchical (Colour-major : Shape-minor dimension)	
10 blue squares 10 blue circles	10 red moons 10 red triangles
(H2) Hierarchical (Shape-major : Colour-minor dimension)	
10 blue squares 10 red squares	10 green circles 10 yellow circles

hierarchical material with colour as the major dimension and shape as the minor dimension. There were two colours and four shapes, each colour being characterised by two shapes and each shape by only one colour.

Every subject underwent a single session in which there were two successive trials, the original trial followed by the shift trial. Each of the three subgroups underwent a different condition. Subgroup A shifted from one dimension to the other with non-hierarchical material. Subgroup B shifted from the minor to the major dimension with hierarchical material. Subgroup C shifted from the major to the minor dimension, also with hierarchical material. The main question of the experiment was whether or not subjects in Subgroups A and B would show the same number of errors in the shift trial, since Subgroup A in the original trial had to place together stimulus aspects on the basis of which they would have to sort in the shift trial, whereas the subjects in Subgroup B would not. The purpose of including Subgroup C was mainly as a control for Subgroup B, since the original trial procedure for Subgroup C was identical to the Shift trial procedure for Subgroup B.

Half the subjects in each subgroup (the colour-shape subjects) sorted by colour in the original trial and by shape in the Shift trial, and the other half (the shape-colour subjects) by shape in the original trial and by colour in the shift trial. The material used by the colour-shape subjects and the shape-colour subjects in each subgroup is given in Table 2. The design which is shown in this

TABLE 2

MATERIAL USED BY THE SUBJECTS IN EACH SUBGROUP.

	Shape-Colour Subjects	Colour-Shape Subjects
Subgroup A	NH	NH
Subgroup B (Minor-Major) ..	H1	H2
Subgroup C (Major-Minor) ..	H2	H1

table was such that it was possible to estimate the degree of interference in the shift trial scores of each subgroup by comparing these with the original trial scores of other subjects.

Subgroup A acted as its own control in that the original trial procedure us with the colour-shape subjects in Subgroup A was identical to the shift trial procedure used with the shape-colour subjects in this Subgroup. Similarly, the original trial procedure for the shape-colour subjects in Subgroup A was identical to the shift trial procedure for the colour-shape subjects in the same subgroup. Thus, for example, in this Subgroup a colour-shape subject would in the original trial have to sort cards into blues and reds, which was the same task as the shape-colour subject had to perform with the same material in the shift trial.

Subgroup C acted as a control for Subgroup B and *vice versa.* This was because, as can be seen from Table 2, the material used by the colour-shape subjects of Subgroup B was the same as the material used by the shape-colour subjects of Subgroup C, and the material used by the shape-colour subjects of Subgroup B was the same as the material used by the colour-shape subjects of Subgroup C. The effect of this was that it was possible to measure the degree of interference in the shift trial scores of the colour-shape subjects of Subgroup B by comparing them to the original trial scores of the shape-colour subjects of Subgroup C, and also to measure the degree of interference in the shift trial score of the shape-colour subjects of Subgroup B by comparing them to the original trial scores of the colour-shape subjects of Subgroup C. In exactly the same way in the original trial scores of Subgroup B acted as controls with which to measure interference in the shift trial scores of Subgroup C.

The procedure was not a free sorting one because in order to ensure that the correct dimension was used in each trial sample cards were used to indicate this dimension. The procedure was the same in both trials. There were four shallow boxes placed on a table in front of the subject. Sample cards were placed in front of the boxes which indicated which colour should be placed in which box during the colour sorting trials and which shapes should be placed in which box during the sahpe trials. In the colour trial, cards of different colours were used as sample cards and these were all of oblong shape. In the shape trial, cards of different shapes were used and these were all black. The experimenter pointed at these sample cards and said " You see these cards, well, I want you to put the cards like this one in this box, etc."

The forty experimental cards were then presented one at a time to the subject in a predetermined random order. Correction procedures were used so that an incorrectly placed card was immediately replaced correctly by the experimenter. The shift trial followed immediately after the original trial and been completed.

In the trials where non-hierarchical material and the major dimension of the hierarchical material were being sorted only two aspects of the relevant dimension were being used and, therefore, there were two sample cards and the subject only sorted into two of the boxes. In the trials where the minor dimension of the hierarchical material was being sorted, there were four aspects utilised, and accordingly, there were four sample cards and four boxes utilised.

III.—RESULTS.

The main aim of the experiment was to compare the shift trial performances of Subgroup A and Subgroup B. Both these subgroups had to sort material into two boxes in the shift trial on the basis either of two colours or of two shapes.

However, since the material used by these two subgroups was different, it was necessary to check whether any differences between the shift trial performance of these two subgroups had been caused by different levels of difficulty in the material which they were using.

It was possible to do this because the tasks undergone by the shape-colour subjects of Subgroups A and B in the shift trial were identical to those undergone by the colour-shape subjects of Subgroups A and C, respectively, in the original trial. Similarly, the tasks undergone by the shape-colour subjects of Subgroups A and B in the shift trial were identical to those undergone by the colour-shape subjects in Subgroups A and C, respectively, in the original trial. Therefore, in order to check the possibility of different levels of difficulty in the tasks undergone by Subgroups A and B in the shift trial, the original trial scores of Subgroups A and C were first compared in an analysis of variance. This analysis, together with the mean error scores is presented in Table 3, where it will be seen that there was apparently no difference between subgroups in either group, but a large difference between the two groups, since Group II, the older children, made less errors. The analysis of the original trial scores of Subgroups A and C yielded a significant Groups term (P ·001). The Materials term which dealt which possible differences between colour and shape sorting was not significant, nor was the Conditions term, which dealt with differences between Subgroups A and C. The lack of a significant Conditions term indicated that the tasks undergone by subgroups A and B in the shift trial were, in fact, of the same level of difficulty as each other and, therefore, that any differences between

TABLE 3

MEAN ERROR SCORES OF SUBGROUPS A AND C IN THE ORIGINAL TRIAL AND ANALYSIS.

Mean Error Scores.

	Colour				Shape			
	Subgroup A		Subgroup C		Subgroup A		Subgroup C	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Group I	2·375	2·176	3·375	2·288	3·500	2·450	3·750	3·345
Group II	1·125	0·060	1·250	1·369	1·000	0·707	0·875	1·053

Analysis.

Source	SS	d.f.	F	P
Groups	76·562	1	17·793	·001
Materials	1·000	1	0·232	N.S.
Conditions	1·562	1	0·363	N.S.
Groups × Materials	4·000	1	0·930	N.S.
Groups × Conditions	1·563	1	0·363	N.S.
Materials × Conditions	1·000	1	0·232	N.S.
Groups × Materials × Conditions	1·813	1	0·421	N.S.

subgroups A and B in their shift trial performance would be due to the effects of their different previous experiences in the original trial.

Accordingly, the shift trial scores of Subgroups A and B were also subjected to an analysis of variance. The analysis, together with the mean error scores, is presented in Table 4. This table shows that there were again differences between groups, the older group making less errors. It also shows that Subgroup A in both groups made more errors than Subgroup B. These differences were subjected to an analysis of variance which yielded a significant Groups Term (P<·001), and a significant Conditions term (P<·01) which was caused by the difference between Subgroup A and Subgroup B. Thus, in the shift trial Subgroup A consistently made more errors than Subgroups B, and the analysis of the original trial scores of Subgroups A and C had shown that this must be due to differences between Subgroups A and B in their previous experience during the original trial.

TABLE 4

MEAN ERROR SCORES OF SUBGROUPS 1 AND 2 IN THE SHIFT TRIAL AND ANALYSIS.

Mean Error Scores and t Tests of Differences with Controls.

	Mean	S.D.	Shape t	
Group I, Subgroup A	10·375	5·678	2·125	p N.S.
Group I, Subgroup B	6·250	2·488	1·425	p N.S.
Group II, Subgroup A	3·75	1·479	4·677	p<·001
Group II, Subgroup B	1·875	1·763	1·289	p N.S
	Mean	S.D.	Colour t	
Group I, Subgroup A	5·625	3·080	2·281	p<·05
Group I, Subgroup B	3·250	3·073	0·108	p N.S.
Group II, Subgroup A	3·125	2·171	2·203	p<.05
Group II, Subgroup B	1·500	1·414	0·370	p N.S.

Analysis.

Source	SS	d.f.	F	P
Groups	232·562	1	23·423	·001
Materials	76·562	1	7·711	·01
Conditions	100·000	1	10·072	·01
Groups × Materials	45·563	1	4·589	·05
Groups × Conditions	9·000	1	0·906	N.S.
Materials × Conditions	4·000	1	0·403	N.S.
Groups × Materials × Conditions	2·250	1	0·227	N.S.

This analysis also yielded a significant Materials term (P<·01) which was caused by the fact that, in the shift trial, the Subgroups A and B subjects of both groups found it harder to sort by shape than by colour. This was an unexpected result since shape and colour sorting were of the same level of difficulty in the original trial performance of Subgroups A and B. In addition, a significant Groups by Materials interaction (P < ·05) showed that this effect was greater with Group I, the younger subjects, than with Group II.

In Table 4 *t* test figures are given with each of the mean error scores. The *t* tests were a further check on the finding that Subgroup A were more hindered in their shift trial performance than Subgroup B. The *t* tests were of differences between each shift trial mean error score and its relevant control. The control for the shift trial score of the colour-shape subjects in Subgroup A was the original trial score of the shape-colour scores of Subgroup A. Similarly, the original trial scores of the colour-shape subjects of Subgroup A served as controls for the shift trial scores of the shape-colour subjects of Subgroup A. In just the same way the shift trial error scores of the colour-shape subjects of Subgroup B were *t* tested with the original trial scores of the Shape-colour subjects of Subgroup C, and the shift trial scores of the shape-colour subjects of Subgroup B were *t* tested with the original trial scores of the colour-shape subjects of Subgroup C. The actual control scores have previously been shown in Table 3.

It can be seen from Table 4 that in no case were the shift trial scores of Subgroup B significantly different from their controls. However, the shift trial error scores of Subgroup A were significantly greater than their controls, in every case except with the colour-shape subjects of Subgroup A, Group I, and here *t* fell just short of significance. These *t* tests were evidence that Subgroup A were subject to interference in their shift trial performance whereas Subgroup B were not.

The shift trial scores of Subgroup C were not included in the above analysis because the subjects of this Subgroup were sorting four aspects into four boxes in this trial, whereas Subgroups A and B were sorting only two. The controls for the shift trial scores of the colour-shape subjects of Subgroup C were the original trial scores of the shape-colour subjects of Subgroup B, and the controls for the shift trial scores of the colour-shape subjects of Subgroup B. The shift trial scores of Subgroup C were *t* tested against their controls, in order to test for interference of previous experience on shift trial performance. These scores, together with the appropriate controls and the *t*s, are given in Table 5, which

TABLE 5

MEAN ERROR SCORES FOR SUBGROUP 3 IN SHIFT TRIAL TOGETHER WITH CONTROLS.

Mean Error Scores.

	Shift Trial		Original Trial		*t* of differences	
	Subgroup C		Subgroup B (controls)		Between Subgroup C and Controls	
Dolour	Mean	S D.	Mean	S.D.		
Group I ..	8·0	4·821	4·875	4·075	1·194	p N.S.
Group II ..	3·625	3·276	3·875	2·649	0·158	p N.S.
Shape						
Group I ..	10·5	3·464	5·25	5·847	2·855	p<·02
Group II ..	4·125	2·804	3·125	2·666	0·684	p N.S.

shows that there was evidence of interference of the shift trial performance of Subgroup C with Group I subjects but not with Group II subjects. In Group I *t* tests yielded a significant difference ($P < \cdot 05$) in the case of subjects sorting by colour, in spite of a fairly large difference between the mean scores of these latter subjects and their controls. In Group II, there appeared to be no differences between the shift trial scores of Subgroup C and their controls, and both *t*s fell far short of significance.

IV.—DISCUSSION.

The main result of this experiment was that, at both age levels, more errors were made in the shift trial by subjects in Subgroup A who had to sort non-hierarchical material than by subjects in Subgroup B, who had to shift from the minor to the major dimension with hierarchical material. This was clear evidence in support of the hypothesis that, in children, failure to shift is caused by persistence of previous learning to avoid specific aspects of the second possible dimension. With Subgroup A the original trial sorting always involved mixing together stimulus aspects which were necessary for the shift trial sorting. However, with Subgroup B the original trial sorting never involved mixing stimulus aspects necessary for the second trial. The fact that Subgroup B consistently performed better in the shift trial than Subgroup A suggests that failures to shift are due to the carry-over of previous learning to avoid specific stimulus aspects rather than to the formation or a general set to concentrate on one dimension and to ignore the others.

There were two additional findings. The first was a difference between the younger and older subjects in the pattern of the shift trial scores of Subgroup C. These subjects had to shift from the major to the minor dimension with hierarchical material. This involved sorting cards in the shift trial on the basis of aspects which had already previously been mixed in the original trial. However, the extent of the previous mixing of the aspects of the dimension which was relevant in the shift trial was not as great with Subgroup C as with Subgroup A. With the colour-shape subjects of Subgroup C, colour was the major dimension and shape the minor dimension, the original trial sorting of colour involved two colour piles, one consisting of two shapes, and the other of the other two shapes. All the four shapes were never placed in one pile. The fact that there was interference in the shift trial with the younger subjects only, when the original trial had involved partial mixing of the aspects relevant to the shift trial sorting suggests the difficulties involved in disregarding previous learning to avoid specific stimulus aspects might be greater at the lower age level.

The second additional finding was that although colour and shape scores were similar in the original trial, more erorrs were made on shape sorting than on colour sorting in the shift trial scores of Subgroups A and B, and that this effect was greater with the younger group. This result might seem to be contradictory to the main finding of the experiment outlined above in that it suggests that previous sorting on the basis of colour led to a set to ignore shape. However, *t* tests consistently showed significant differences between the shift trial scores of Subgroup A and their controls and no significant differences between those of Subgroup B and their controls, both in the case of colour and of shape sorting. This finding indicated that although shape was more difficult than colour for both groups in the shift trial, the primary cause of failure to sort on the basis of the relevant dimension in the shift trial was previous experience of learning to avoid specific aspects of that dimension in the original trial.

The results of this experiment are in agreement with studies by Kendler and Kendler (1962) of discrimination learning and reversal in children of different age levels. These studies showed that children below 6 are able to reverse dimensions more easily than to reverse the significance of two stimuli within a dimension. They also showed that at higher age levels the opposite pattern occurred. If these discrimination study findings are applied to colour-form sorting test scores, it could be argued that they indicate that failure in adults to shift is caused by a failure to move from one dimension to another, rather than a failure to reverse the significance of specific stimuli. This could mean that failures to shift are the results of different processes in adults than in children of 5 to 7 years.

Acknowledgements.—This research was carried out at the Institut des Sciences de l'Education, Geneva, when the author was the holder of a Mental Health Research Fund Travelling Fellowship. The author is grateful to the Members of the Institut and in particular to Professor B. Inhelder, for the opportunities which they gave him to discuss this work with them.

V.—References.

Goldstein, K., anad Scheerer, M. (1941). Abstract and concrete behaviour : an experimental study with special tests. *Psychol. Monogr.* 53.

Kendler, H. H., and Kendler, T. S. (1962). Horizontal and vertical processes in problem solving. *Psychol. Rev.*, 69, 1-16.

Inhelder, B., and Piaget, J. (1964). *The Early Growth of Logic in the Child.* London : Routledge and Kegan Paul.

Reichard, S., Schneider, M., and Rapaport, D. (1944). The development of concept formation in children. *Amer. J. Orthopsychiatr.*, 14, 155-161.

Thompson ,J. (1941). The ability of children of different grade levels to generalise sorting tests. *J. Psychol.*, 11, 119-126.

(*Manuscript received* 23*rd August*, 1966)

PREMATURE RETIREMENT OF WOMEN TEACHERS FROM GIRLS' AND MIXED SECONDARY SCHOOLS

By R. R. DALE
(*University College of Swansea*)

SUMMARY. This article attempts to inject objective evidence into the arguments about co-education. The rates of premature retirements of women teachers in mixed and girls' secondary schools are compared as rough indices of the level of nervous tension in such schools. The analysis includes data on the age distribution of women teachers, based on a national sample. As a by-product is found an interesting difference in the rate of retirement from grammar and from secondary modern schools.

I.—INTRODUCTION.

THE long continued debate on the rival merits of co-education and segregation in our schools has been sustained by many theoretical arguments, but supported by very few facts. An attempt has previously been made to supply some of these facts by assessing the experience of teachers, especially the experience of those who have taught both in mixed and single-sex grammar schools (Dale, 1955, 1965). It must be said, however, that although this evidence is weighty, it is not necessarily conclusive. Human nature is such that even the verdict of experience may be biased by the climate of social opinion. Moreover, human judgment, even if collective, is not *ipso facto* infallible. Researchers, therefore, should endeavour to provide evidence which is more objective, and the article seeks to point out ways in which this might be obtained. Although in itself, it examines only one facet of the total situation, and that not conclusively, it seeks also to act as a signpost to research. In such a virgin territory it is unlikely that any contribution will be at once definitive; the map cannot be drawn immediately, rather do we first search for good triangulation points.

One of the arguments advanced in favour of mixed schools is that in a girl's school the atmosphere is liable to become too tense, that head mistresses tend to have a pettifogging over-insistence on little details, wearing thin the nerves of their assistants; that mistresses, deprived of the moderating influence of assistant masters, tend to become over-conscientious, with detrimental effect on their health; that in a mixed school the presence of both sexes creates an atmosphere which is less strained, with better relations among staff and between pupils and staff, thus producing a happier—healthier—environment.

The counter argument runs as follows: In mixed schools there are often feuds between men and women staff, producing nervous tension which may have a bad effect on health. Again, some women cannot control boys, and the indiscipline is liable to produce emotional upsets; in any case, the long continued tension is inimical to health. Finally, women in a mixed school are placed in a position of inferiority; they can rarely become heads of mixed schools and some become frustrated; this prolonged frustration may have a deleterious effect on physical and mental health.

The importance of such tensions, if they exist, needs little stressing. Adults who are anxious and disturbed cannot have a good influence on the personalities of children. Individuals function best where the atmosphere is as free as possible of unpleasant tensions. A minimising of strain would improve the health and personality of the teachers, resulting in the better emotional development of the pupils. There is no suggestion that the sex of the school is the only factor in the situation, but merely an intention to examine whether it is one appreciable source of differentiation.

II.—PROCEDURE.

These arguments for and against co-education would chase each other round for ever if research did not produce new evidence. The problem was to obtain this evidence from a population sophisticated to research projects, sensitive to enquiries on this particular subject, and often loyal to individual schools. The direct approach by paper and pencil tests of neuroticism, etc., would have been rather naive in this context, and especially so if the teachers had any inkling that the test was connected with research on co-education.

It was, therefore, decided to proceed by means of facts which would be less open to dispute than would paper and pencil tests. The criterion used was the health, especially mental health, of women teachers in the two types of school, measured by the incidence of premature retirement on breakdown pension, and the age at which such retirement occurred. The question immediately arises, " Why should physical health be included as well as mental health?"

There is a growing awareness that emotional strain frequently, if not always, has a profound effect on physical health. This is now well documented, the range of illnesses including duodenal ulcers, arterial hypertension, allergies, many skin ailments, rheumatoid arthritis and even diabetes. This being so, the inclusion of all breakdown retirements, both mental and physical, is justifiable. It is also assumed that ' purely ' physical illness will occur in approximately the same frequency, for teachers of the same age, among women teachers in girls' schools as in mixed schools.* As this assumption may be affected by differences between these types of school in their urban-rural sitings, this question will presently receive examination. However, in order to meet possible criticism about combining the physical with the mental, a further analysis is made in which those illnesses classed as ' purely physical,' as determined by two doctors, are excluded from consideration. The Ministry's medical staff were also requested to distinguish those illnesses which were mainly nervous.

The collection of the information proved unexpectedly difficult. A scheme was devised by which the then Ministry of Education collected the information year by year (from 1950 to 1956) as each case of retirement on breakdown pension was sanctioned. The rate of retirement, however, became lower than expected, somewhat inexplicably even to the Ministry, so that the total number of cases from mixed grammar schools was fewer than was desirable statistically. The Ministry by the Spring of 1957, was unable to give further assistance. It was, therefore, decided to increase the number of cases by including secondary modern schools—and incidentally comprehensive and multilateral schools—as the raw material was already available, and it was thought that if there was any pronounced difference between breakdown retirements from mixed and girls' grammar schools, there would be a similar (though not necessarily the same) difference in mixed and girls' secondary modern schools. This theory would be checked by analysing the grammar and modern schools separately as well as in combination. For both grammar and modern schools the investigation was confined to State schools.

---

* The same assumption, under the null hypothesis, is also made about mental illness, and the possible operation of the same urban-rural variable is borne in mind.

The data included the age of the teacher, the names and type of her schools for the last five years of service, marital status, last date of paid service, date from which the pension was paid, whether head or assistant, and the cause of retirement. For grammar schools the list included all teachers whose breakdown pensions had been granted during the years in question, irrespective of the last date of paid service. The secondary modern school list comprised only those teachers whose last date of paid service was on, or later than, the 31st December, 1949. The criterion was one year's service in the school from which they retired, or in a previous school of the same type, and same sex of pupil. In fact, no case was included where the teacher had served in both mixed and girls' schools during the last two years of teaching ; almost all teachers had been at least five years in the same type of school. The operative date was taken as the last date of paid service.

III.—RESULTS.

The incidence of premature retirement in maintained and transitionally assisted grammar schools, within the objective limitations laid down by the experiment, is given in Table 1.

TABLE 1

WOMEN TEACHERS : RETIREMENT ON BREAKDOWN PENSIONS.†

Year	From Girls' Schools	Teachers in Girls' Schools	Incidence	From Mixed Schools	Teachers in Mixed Schools	Incidence
1950	16	8,813	1 in 550	3	3,265	1 in 1,088
1951	11	8,968	1 in 815	6	3,231	1 in 538
1952	19	9,078	1 in 478	5	3,279	1 in 656
1953	19	9,189	1 in 484	4	3,246	1 in 811
1954	20	9,330	1 in 466	5	3,256	1 in 651
1955	10	9,452	1 in 945	6	3,330	1 in 555
1956	7	9,664	(1 in 1,380)	2	3,422	(1 in 1,711)
Total ..	102		(1 in 632)	31		(1 in 743)

N.B.—Two of the thirty-one from mixed schools and two of the 102 from girls' schools did not serve five years in the type of school from which they retired. 1956 was an incomplete year.

† Note the important comment later in the text, discussing the population eligible for this pension.

Although there are large fluctuations, due to the smallness of the number of retirements, the mixed schools have a superior record in five years and the girls' schools in two. To take this, however, at its face value would be misleading, as will be seen later, so that no statistical tests will be applied. If we take as a basis the average number of teachers in the schools over the seven years, the percentage of breakdown retirements over the whole seven-year period is 1·1 per cent. for girls' schools (or 1 in 90) and 0·94 per cent. (or 1 in 105) for mixed schools. (This raw difference is not statistically significant.) If we use instead the criterion of the average age of retirement, that from mixed schools was 53 years 7·8 months for the thirty-one cases over seven years, and that for the girls' schools was 52 years 3·0 months for 102 cases. (Statistically a non-significant difference, the critical ratio being 1·38 ; the standard error is comparatively large because of the small number of cases from the mixed schools.) In six years out of seven the average breakdown retirement age was higher in the mixed schools than in the girls'. The median retirement age for the mixed schools was 54·75 years, and 53·75 years for the girls' schools.

A comparison of the two age distributions of premature retirements, is given in Table 2, partly in case a future researcher might wish to add to its evidence. Although it is interesting that there is no retirement from mixed schools—as defined by our objective criteria—below the age of 40, while eight cases or 7·9 per cent. of the girls' schools total retire before that age, no definite conclusions can be drawn, as the number retiring from mixed schools is so small, and the fluctuations of percentages for particular age groups are likely to be large.

TABLE 2

COMPARATIVE AGE DISTRIBUTIONS OF PREMATURE RETIREMENTS.

	Mixed Schools		Girls' Schools	
Age group	No.	Percentage	No.	Percentage
30—34	0	0	1	(c1·0)
35—39	0	0	7	6·9
40—44	2	(6·5)	8	7·8
45—49	3	(9·7)	9	8·8
50—54	11	35·5	34	33·3
55—59	15	48·4	42	41·2
60—64	0	0	1	(c1·0)
Totals	31	100	102	100

An attempt was made to allow for the difference in the urban-rural situation of the mixed and girls' grammar schools by noting which retirements were from Rural District Areas, with a view to excluding them (and the relevant basic populations) from the statistics. Unfortunately, this was an ineffective division, yielding ' rural ' numbers which were much too small. Their omission merely reduced the number retiring from girls' schools to 100 and from mixed schools to twenty-nine and lowered the average age of retirement very slightly to 53 years 7·3 months for mixed schools and 52 years 2·6 months for girls schools. As their omission would have made so little difference, they were retained. The fact that many of the girls' schools but few of the mixed schools were in the large cities, must therefore be taken into account in the interpretation of the statistics.

The inclusion of the few comprehensive, bilateral and multilateral schools added only two cases to the total for girls' schools and one case to the mixed, reducing the average age of premature retirement in girls' schools to 52 years 2·9 months, and in mixed schools raising it slightly to 53 years, 9·6 months.

During the assessment of the data it was seen that another unwanted variable might be affecting the results, namely, a difference in the age distribution of teachers in mixed compared with girls' grammar schools. As no published statistics on this point existed and neither the Ministry nor the local education authorities could supply the information, it was decided that the data were so crucial to the enquiry that a survey of the age distribution of women teachers would have to be made, even though this would show the distribution only for the year after the retirement survey ended. A national sample of 2,688 teachers, stratified into county, county borough and London areas, was drawn.

The results given in Table 3 show an appreciable difference in the age distribution of women teachers between the two types of school. The figures are in percentages, with standard errors in italics. Relevant Bilateral and Comprehensive schools are included, though these were few.

TABLE 3

AGE DISTRIBUTION OF WOMEN TEACHERS IN MAINTAINED GRAMMAR SCHOOLS
(31st August, 1957)

The sample comprised 2,688 teachers.

	County		County Boroughs		London		Totals		
Age Group	Mixed	Girls'	Mixed	Girls'	Mixed	Girls'	Mixed	Girls'	Combined
20—29	45·62	37·95	48·78	36·14	41·18	33·5	45·8	36·7	39·47
S.E.	*2·62*	*1·75*	*2·06*	*1·48*	*3·6*	*1·78*	*1·71*	*0·98*	*0·86*
30—39	18·43	21·58	20·21	20·12	18·63	22·45	18·65	21·1	20·34
S.E.	*2·04*	*1·48*	*1·66*	*1 24*	*2·85*	*1·57*	*1·34*	*0·83*	*0·71*
40—49	18·43	19·05	12·54	21·05	18·63	20·66	17·73	20·0	19·34
S.E.	*2·04*	*1·41*	*1·37*	*1·26*	*2·85*	*1·53*	*1·31*	*0·8*	*0·69*
50 plus	17·52	21·43	18·47	22·69	21·57	23·36	17·83	22·15	20·84
S.E.	*2·00*	*1·48*	*1·63*	*1·29*	*3·0*	*1·6*	*1·3*	*0·85*	*0·71*
Total	100	100	100	100	100	100	100	100	100

The table shows that there is a persistent common pattern. In each of the strata, county, county borough and London, the mixed schools have a decidedly greater proportion of teachers under the age of 30. Correspondingly, in both the age groups 50 plus and 40 to 49 the girls' schools have a slightly higher percentage of teachers. The differences in the total percentages, between mixed and girls' schools, though small in all but the 20 to 29 age group, are therefore, a reflection of a consistent trend in all strata. The differences between girls' and mixed schools in the percentage of teachers over 40 is statistically highly significant. Similarly for teachers aged 50 plus. (Between the three strata themselves there is a surprisingly small difference.)

This unexpected difference in the age distribution of women teachers in girls' grammar schools as compared with mixed, though not large, necessitated a re-appraisal of the statistics. Most of the premature retirements are of teachers aged 40 and over, and of these by far the greater proportion are of teachers aged 50 and more. As the percentage of teachers aged 40 plus is rather higher in girls' schools than in mixed, it could be argued that the difference in age distribution is at least one cause of the rather lower incidence of premature retirements in mixed schools—if one concedes that this lower incidence exists. On the other hand, it should be noted that with the second criterion—the average age—the age distribution factor acts in the other direction, i.e., towards producing an average age of premature retirement higher in girls' schools than in mixed, whereas in fact the opposite appears to be the case.

The operation of this factor will clearly be seen in an extreme example; if all teachers in girls' schools were more than 55 years of age, the average age of premature retirement could not be lower than 55, and if all teachers in mixed schools were below 50, the average age of premature retirement for these schools must be 50 or lower. Though the difference in the retiring age from mixed and girls' schools is a little short of conventional statistical significance, its existence, in spite of the unfavourable age distribution in the mixed schools, provides a pointer for future research, though the urban-rural variable must be borne in mind.

The position is further complicated. The proportion of the teacher population which had not completed the necessary ten years of contributory service to qualify for a breakdown pension was, in 1957, greater in mixed than in girls' schools. Though we haven't the data for previous years, it is highly likely that the above difference would already have been in existence and that it would be appreciable, though its precise amount might have been different. Under these circumstances a rough calculation was made of the incidence of premature retirements, based on the average number of teachers eligible for this pension over the seven years. The rate for mixed schools at 1 in 57·6 was almost identical to that for girls' schools at 1 in 57·2. The virtual equality of these rates does not, of course, prove that there is equality in practice, because we do not know the age distribution of the teachers in 1950, at the start of this survey. What we can say is that we have found no startling difference between the two types of school in the incidence of premature retirement, and, in view of the smallness of the numbers and the uncertainty about the precise age distribution in early years,only a large difference would have been dependable.

The examination of the comparative *age* of retirement on breakdown pension is also affected by the difference in the proportions eligible for it in the two populations. When we exclude the teachers who are under 30 years of age, the proportions in the remaining age ranges are very similar for both mixed and girls' schools, the largest difference being a 2 per cent. greater proportion of teachers aged 50 plus in girls' schools than in mixed (35·02 per cent. compared with 32·89 per cent.). Though this 2 per cent. difference would have a small influence in increasing the average age of premature retirement in girls' schools and, therefore, in hiding the 'true' difference between girls' and mixed schools and in pushing this difference towards statistical significance, the result would not be decisive enough to justify further calculation. None the less, this age difference in favour of mixed schools, though statistically not reliable, is interesting, and may encourage further enquiry on a larger scale. In any such research an effort would have to be made to equate the two groups on the urban-rural variable.

Another approach to the problem is via mental ill-health, both in proportion to the physical ill-health of the teachers and also in proportion to the number of women teachers in the two types of school. One of the doctors at the Department of Education and Science was, therefore, asked to classify the cases into 'Mental' and 'Physical.' This process allocated 22·5 per cent. of premature retirements from girls' schools to the 'Mental' category (23 out of 102), and 25·8 per cent. from mixed schools (8 out of 31). The mixed school percentage, because of the small number on which it is based, is rather unstable, though one must bear in mind that the cases are really serious ones that have been collected over a period of seven years. As a double check, the list of cases (without names and without indication of type of school) was given to two other doctors to classify independently of each other. In this instance, an intermediate category (B) was inserted. Both doctors placed 35·5 per cent. of cases from mixed schools in the Mental (A) and Partly Mental (B) categories, and were in close agreement about those from girls' schools, the figures being 36·3 and 38·2 per cent. If we take the mental category alone the outside doctors agree at 23·5 per cent. for girls' schools (very close to the Ministry's own figures), but disagree over several cases for the mixed school, one doctor placing only six (19·4 per cent.) in this category, and the other placing ten in the same category (32·3 per cent.). The Ministry's figure is the average between the two extremes. The differences between the mixed and girls' schools is so small that there would be no point in pursuing this approach any further.

The last method used was the rather intriguing one of examining the number of teachers in the two types of schools who retired ' in pairs ' or ' triples' from the same school over the seven-year period. The use of the word ' pairs needs qualifying : it is not inferred that the teachers retired at exactly the same time but that two or more teachers retired on breakdown pension from one school during the seven years. At first sight, the results from mixed and girls' schools appear to show an interesting difference. From the girls' schools nineteen teachers retired in pairs or triples from nine schools, out of an overall total of 102 cases from 402 schools and 9,230 teachers. From the mixed schools four teachers retired from two schools, out of an overall total of thirty-one teachers from 372 schools and 3,290 teachers. In order to discover whether these distributions were very different from what might be expected by chance, use was made of the Poisson distribution formula. The results are in Table 4.

TABLE 4

DISTRIBUTION OF WOMEN TEACHERS RETIRING FROM GRAMMAR SCHOOLS.

	Girls' Schools Schools 402 : Teachers at ' risk,' 5,843		Mixed Schools Schools 372 : Teachers at ' risk,' 1,773	
	Actual	By chance (Poisson)	Actual	By chance (Poisson)
Single cases	83	80	27	28·4
Two (or more) from same school	19	22	4	2·6
Total	102	102	31	31

It will be seen that the apparent greater tendency of the teachers from girls' schools to retire ' in pairs ' from the same school is illusory. The smallness of the mixed school figures also prevents us from making any reliable statistical comparison of the two types of school.

The most important weakness in the survey up to this point has been the smallness of the number of teachers retiring from the mixed schools. An endeavour was, therefore, made to minimise the weakness by adding to the survey those teachers who retired prematurely from secondary modern schools. Teachers who spent less than one year in their last school were omitted unless their previous experience (up to one year) had been in the same type of school. The ninety-two teachers retiring on breakdown pensions from girls' secondary modern schools had an average of 52 years 7·4 months, which was very similar to the 52 years 8·9 months of the seventy-two retiring from mixed secondary modern schools. When the very small number of retirements from schools in rural district council areas were removed, each average dropped slightly, the mixed to 52 years 6·2 months and the girls' schools to 52 years 6·7 months. The *incidence* of retirement was also close in the two cases, being one in 141 for the girls' schools and one in 147 for the mixed, over the seven years, based on the average number of teachers. With the removal of schools situated in rural district council areas, the figures became almost identical at one in 140 for girls schools and one in 137 for mixed. Once again, however, there were appreciable differences in the age distributions of teachers in the two types of school, as shown in Table 5, estimated from a national stratified sample of 2,087 teachers.

TABLE 5

SECONDARY MODERN SCHOOLS. AGE DISTRIBUTION OF WOMEN TEACHERS AT 31ST AUGUST, 1957.

(The figures are in percentages, with standard errors in italics.)

The sample comprised 2,087 teachers.

	County		County Boroughs		London		Totals		
Age Group	Mixed	Girls'	Mixed	Girls'	Mixed	Girls'	Mixed	Girls'	Com-bined
20—29	44·4	38·2	51·46	46·69	49·3	48·9	46·27	42·44	44·39
S.E.	*2·83*	*2·09*	*2·88*	*2·21*	*2·89*	*2·42*	*1·71*	*1·31*	*1·04*
30—39	17·51	17·03	20·07	13·65	18·1	11·43	18·09	15·23	16·69
S.E.	*2·17*	*1·62*	*2·32*	*1·52*	*2·23*	*1·54*	*1·33*	*0·95*	*0·78*
40—49	20·54	21·78	17·15	18·55	10·1	14·6	19·16	19·91	19·53
S.E.	*2·31*	*1·78*	*2·18*	*1·72*	*1·74*	*1·71*	*1·35*	*1·06*	*0·83*
50 plus	17·51	22·97	11·31	21·1	22·5	25·1	16·48	22·41	19·39
S.E.	*2·17*	*1·81*	*1·83*	*1·81*	*2·41*	*2·1*	*1·27*	*1·11*	*0·83*
Total	100	100	100	100	100	100	100	100	100

The totals columns show that the teachers in girls' secondary modern schools are somewhat older than those in mixed secondary modern schools. The difference between the proportions over 40 years old is statistically highly significant. This difference in age ought to tend to produce a slightly higher rate of premature retirement in the girls' schools (whereas there was little difference when the rural schools were removed) and a higher average age of retirement (whereas the two average ages were in fairly close agreement). The closeness of both the incidence and the average age of premature retirement from girls' and from mixed schools, combined with the unertainty about the respective age distributions in 1950, and the urban-rural variable, makes further analysis on these lines unprofitable. For the same reasons no exhaustive analysis is made of the combined grammar, bilateral, comprehensive and secondary modern figures. It is reported, however, that for these combined schools, using the same basis as for the other schools, the incidence of breakdown was 1 in 65·94 for girls' schools (1·52 per cent.) and 1 in 74·86 (1·34 per cent.) for mixed schools.

We therefore turn to the next criterion, the incidence of mental ill-health. In the case of the girls' schools the official doctors placed twenty-four out of ninety-two (26·1 per cent.) retirements in the 'mental' category, and for the mixed schools twenty-two out of seventy-two or 30·6 per cent. These were judgments based on the short official diagnostic descriptions of the illness. Two private doctors acting independently of each other graded the cases into three classes by adding an ' Intermediate ' category. They used the same evidence as the official doctor used when he placed the cases into categories, and they did not know whether the individuals came from mixed or girls' schools. Their judgment tends to reverse the slight difference in favour of girls' schools which resulted from the official doctor's classification. If the ' Mental ' and ' Partly Mental ' (or Intermediate) classes are combined, the first doctor placed in this category 30·2 per cent. of the retirements from girls' schools and 27·5 per cent. of the retirements from mixed schools. The corresponding figures for the second doctor were 38 per cent. (girls' schools) and 33·3 per cent. (mixed). Separating the two classes we find that the first doctor placed exactly 25 per cent. of the cases in the ' Mental ' class for both girls' and mixed schools, and the second gave 21·7 per cent. for the girls' schools and 23·6 per cent. for the mixed. The results are too fluctuating and the gaps are too narrow to justify any attempt to ascertain their statistical dependability.

The remaining criterion—that of retirement in ' pairs ' —was then examined. From girls' secondary modern schools twenty teachers retired prematurely (in ' pairs ') from ten schools, out of an overall total of ninety-two cases from 872 girls' schools (over 7 years), and 7,462 teachers at risk (average for 7 years). Turning to mixed secondary modern schools, two teachers retired prematurely (in ' pairs ') from one school, out of an overall total of seventy-two cases from 1,692 schools and 5,695 teachers at risk (average). These figures were then compared with those expected by chance. The result is given in Table 6.

TABLE 6

DISTRIBUTION OF WOMEN TEACHERS RETIRING FROM SECONDARY MODERN SCHOOLS.

	Girls' Schools Schools 872 Teachers at ' risk,' 7,462		Mixed Schools Schools 1,692 Teachers at ' risk ' 5,695	
	Actual	By chance (Poisson)	Actual	By chance (Poisson)
Single cases	72	82·8	70	69·1
Double or triple cases (Individuals)	20	9·2	2	2·9
Total	92	92	72	72

The numbers for the mixed schools are too small for any reliable comparison to be made with the girls' schools. In the case of the girls' schools themselves, the difference between the actual retirements (in ' pairs ') and the chance occurrences is much greater, and is statistically significant beyond the 0·025 level. This is in line with the argument that some girls' schools tend to have a high level of tension among the staff, but as the analysis of the data has uncovered no consistent pattern of this kind, it could by no means be said that this relationship had been proved. As evidence against this argument could be placed the apparent lack of large differences between mixed and girls' schools in the incidence of premature retirement. Another explanation could be that, in spite of the statistical significance, this may be a chance effect—

though the odds against it are high. A reasonable conclusion would be that an interesting relationship may have been uncovered which invites further research on its stability and causation.

One of the most interesting aspects of the work is an incidental finding, not related to the comparison between mixed and girls' schools. In Tables 4 and 6 the rates of breakdown for teachers eligible for this pension are some 40 per cent. higher in grammar than in modern schools, and this holds good for both mixed and single-sex schools. An inspection of the age distribution tables shows that this factor cannot be held responsible for any of the differences. A close examination of the enquiry data has produced no difference in procedure or criterion which could account for this lesser incidence in modern schools. None the less, confirmative evidence from an independent study over a different period of years would be needed before one could claim that the difference is established.

IV.—CONCLUSIONS.

This reseach has attempted to introduce more objective criteria into the study of the comparative effects of co-educational and single-sex schooling. If it had established the existence of an appreciable difference between the breakdown rates in mixed and single-sex schools it would have been hailed as a major break-through in this field ,starting off a new line of enquiry. That it has not done this reduces its value, but it is hoped other researchers will be able to benefit from this pioneer work. Its weaknesses—the smallness in the number in the mixed schools samples, the absence of information about the age distribution of teachers in mixed and girls' schools at the commencement of the survey, and the lack of data about the urban-rural variables—could be overcome in future research though they were difficult to foresee in the planning of this study. Another possible source of error is the nature of the intake of teachers into mixed and girls' schools ; there may be tendency for the two samples to have—on the average—somewhat different personality patterns, and a differing liability to illness. Other suggestions for research are the extension of the objective criteria to include measures like the incidence of illness among the staff, pupil truancy, juvenile delinquency, pupil maladjustment, etc.

In a project of this nature, only the discovery of large differences between mixed and girls' schools could give a definite result. No such consistent pattern of difference has been found. Though there is a slight tendency for the mixed schools to have a rather better health record among the staffs, as judged by the breakdown criterion, and though the chance explanation is not likely—it cannot be dismissed—yet the predominantly urban siting of the single-sex schools probably places them at a disadvantage and this might be an adequate explanation of their less good record.

ACKNOWLEDGEMENTS.—The writer is heavily indebted to the then Ministry of Education, and in particular to Mr. D. S. Hutton, of the Statistics Branch, for the facilities they offered and their kindness and forbearance during the work. Mr. G. F. Peaker, H M.I., and Mr. R. Startup are thanked for their comments, particularly on the use of the Poisson distribution, Mrs. R. Lewis for her very efficient secretarial assistance, and Dr. Jean James and Dr. P. James for their kind co-operation.

V.—REFERENCES.

DALE, R. R. (1955). Co-education : the verdict of experience. *Brit. J. Educ. Psychol.*, 25, 145-157.

DALE, R. R. (1965). Co-education : the verdict of experience II—a qualitative approach. *Brit. J. Educ. Psychol.*, 35, 195-209.

(*Manuscript received* 23*rd August*, 1966)

AN EXPERIMENT ON MULTIPLE-CHOICE *VERSUS* OPEN-ENDED ANSWERING IN A VOCABULARY TEST

BY A. W. HEIM AND K. P. WATTS
(Psychological Laboratory, University of Cambridge)

SUMMARY. An experiment was conducted in which the results of multiple-choice and open-ended techniques were compared on a test of vocabulary (Test E of the Self-judging Vocabulary Scale). The subjects were two groups, each consisting of fifty naval ratings, equated on intelligence test AH 4. Group M took the vocabulary test in its multiple-choice form immediately followed by the same test in open-ended form ; for Group O, the order was reversed.

Significant differences were found between multiple-choice and open-ended scores, the former being substantially higher than the latter. The order in which the two forms of vocabulary test were taken was found to influence (*a*) the subjects' responses and (*b*) the association between intelligence test and vocabulary test scores. On both forms of the vocabulary test, the subjects' scores proved to be closely related to their self-estimates of word-familiarity.

I.—INTRODUCTION.

THE experiment was designed to throw light on the following questions with reference, particularly, to vocabulary testing. (*a*) To what extent does the multiple-choice technique provide an easier task for the subject than open-ended answering ? (*b*) What is the effect on the subject's responses, of taking the same test first in multiple-choice form and then in open-ended form ? And conversely ? (*c*) How is the subject's own estimate of his relevant word-knowledge related to the correctness of his responses to the test items ? (*d*) Is the association between intelligence test score and vocabulary test score influenced by the form in which the vocabulary test is taken ?

We are particularly interested in question (*a*), owing to the belief (and finding) of Raven (1958) that the difference between multiple-choice and open-ended answering in his Mill Hill Vocabulary Test was extremely small—usually a difference of only one or two points, when used with adults. We suspected that, especially in a vocabulary test, multiple-choice is, in fact, considerably easier than open-ended answering, provided that the multiple choices offered are themselves well understood by the subject.

II.—DESIGN OF EXPERIMENT.

The subjects consisted of two groups, M(=multiple-choice, first) and O (=open-ended, first), each comprising fifty Naval ratings. These 100 sailors had been selected from a larger sample (n=135), since we wished to use groups whose scores were matched on an intelligence test, for our experimental subjects. The vast majority of the sailors were 17–26 years of age ; most of them were from secondary modern schools, though a few had been to grammar schools (roughly equal numbers of the latter in Group M and Group O, respectively) ; they were all junior ratings, mainly seamen, stokers, electricians and naval airmen. The mean ages of Groups M and O were, respectively, 20·7 and 20·8 years.

The 135 sailors took intelligence test AH 4 (1955) and, on the basis of their scores, the two groups of fifty were selected. The mean AH 4 score and standard deviation of Group M were, respectively, 77·9 and 11·15; those of Group O we e 77·8 and 11·13.

Within a fortnight of their intelligence test they had their second session. At this session, the subjects were presented (i) with a list of the forty words which make up the self-judging Vocabulary Test E (Heim and Watts, 1961) and, as is customary with tests of the Self-judging Scale, they were asked to mark each word on the list as *a*, *b* or *c*. They were told that *a* signified that they knew the word in question and could explain its meaning to someone unfamiliar with it; that *b* implied doubt; and that *c* signified that they had never seen the word before and had absolutely no idea what it meant.

(ii) The subjects were then given a second list of words (this time fifty), which they were asked to *a*/*b*/*c* in a similar way. The main purpose of this was to decrease the possibility of subjects' recognising which words they had marked as *a*, etc., when taking the subsequent vocabulary test. This procedure formed part, too, of a different experiment (on the Word-in-context test), (Cook, Heim and Watts, 1963) which will not be dealt with here.

(iii) Having marked the two lists of words, members of the M group were given the vocabulary test, first in its multiple-choice form and, immediately afterwards, in its open-ended form. Members of the O group, however, were given the vocabulary test, first in its open-ended form and then, immediately afterwards, in its multiple-choice form.

For administrative reasons, the testing was conducted over a period of some months. The subjects were tested in sub-groups of six, every member of any one sub-group taking the pair of vocabulary tests in the same order. The test words consisted mainly of nouns and adjectives; there were some which had a somewhat scientific bias (such as ' genetics ' and ' alkaline ') and others which had, rather, a flavour of the humanities (such as ' sinecure ' and ' gullible '). The stimulus words were, deliberately on our part, rather difficult for our subjects but all the multiple-choice responses were well within their scope. In both parts of the test, subjects were asked to guess if they did not know the answer. Naturally, enough, they found it easier to hazard a guess in the multiple-choice than in the open-ended situation. They did, in fact, answer every multiple-choice item but, despite the instructions, they omitted many items in the open-ended form. (Thus, there is a sense in which the open-ended test in effect comprised fewer questions than did the multiple-choice test.)

In the multiple-choice form, the subject scored one for choosing the correct response from the proffered six and nought if he chose any one of the other five. In the open-ended form, however, classifying as merely right/wrong would have led to over-simplication of the issue, since varying degrees of correctness were found among the multifarious answers given. In the interests of consistency and time-saving, we compromised here and elected to score one for a good answer, half for an answer that was neither wholly right nor wholly wrong, and nought for an utterly mistaken answer.

Let us illustrate this system of scoring, with examples drawn from the answers offered for the word ' genetic ': scoring one, ' heredity,' ' the study of genes '; scoring half, ' theories ' ' medical term ' ' a type of mathematics '; scoring nought, ' a plastic object,' ' to do with the Gospels,' ' things peculiar to certain things.' These eight examples are chosen from Group O's answers. It was found that Group M often gave, in the open-ended situation, phrases which they had seen against the stimulus word, in the multiple-choice form of the test. These results are dealt with in Section III C.

III.—RESULTS.

A.—*Relative ease of the two forms of test.*

Table 1 shows the mean scores and standard deviations of Groups M and O on the multiple-choice and open-ended forms of the test.

TABLE 1

COMPARISON BETWEEN MEAN TEST SCORES FOR THE TWO GROUPS.

	Group M (Multiple-choice first) n=50		Group O (Open-ended first) n=50		Groups M+O n=100	
	Mean	s.d.	Mean	s.d.	Mean	s.d.
Multiple-choice score . . .	26·66	5·19	26·16	5·36	26·41	5·28
Open-ended score	19·67	6·91	16·51	6·28	18·09	6·79

The difference between the mean multiple-choice and open-ended scores are significant in each case, $p<0.01$.

The difference between the mean open-ended scores for the two groups is also significant, $p=0{\cdot}02$.

In every case, the difference between the multiple-choice mean score and the open-ended mean score is significant ($p<0{\cdot}01$). It is thus clear that: (*a*) Multiple-choice provides a far easier task than does open-ended answering, when the identical stimulus words are used and the multiple-choice phrases offered are intelligible to the subject. (*b*) The greatest difference is found, predictably, among the subjects who took the test in its open-ended form, *first.*

It is interesting that the fact that Group M took the multiple-choice first effectively reduced the difference, despite the subjects' having had no knowledge of results on their multiple-choice performance. Furthermore, it may be seen that the multiple-choice mean scores of Group M and Group O are almost identical, whereas the two open-ended mean scores differ by just over three points, the difference being significant ($p=0{\cdot}02$). Throughout, the standard deviation on the open-ended form is greater than that on the multiple-choice form.

Table 2 analyses the above findings in greater detail.

TABLE 2

ANALYSIS OF SCORING IN THE OPEN-ENDED TEST FOR THE TWO GROUPS.

	Group M (n=50)	Group O (n=50)
Mean number scoring one in open-ended test	17·4	13·34
Mean number scoring half in open-ended test	4·58	6·36
Mean number scoring nought in open-ended test	9·6	10·08
Mean number omitted in open-ended test	8·42	10·22

It may be seen that the better performance of Group M's members on the open-ended form of the test is due to their relatively high number of ones, although the difference is not significant. They actually score fewer halves than do the members of Group O, though this difference, too, does not reach significance level. It is likely that the subject who has some slight familiarity with the stimulus word will score half in the O situation whereas, in the M situation, he will have chosen correctly on the multiple-choice test, thus enabling himself to score one, when he reaches the open-ended test.

B.—*Effect of order in which vocabulary tests are taken.*

Table 3 shows the mean number of open-ended answers, given by the two groups, using wording identical or similar to that offered in the multiple-choice form, correct and incorrect.

TABLE 3

CHOICE OF WORDING IN THE OPEN-ENDED TEST.

	Group M (n=50)	Group O (n=50)	Difference (M—O)
Mean number using ***identical*** and correct multiple-choice wording in open-ended test	8·4	2·2	6·2 $p<0{\cdot}01$
Mean number using ***similar*** and correct multiple-choice wording in open-ended test	11·2	4·0	7·2 $p<0{\cdot}01$
Mean number using incorrect multiple-choice wording in open-ended test (identical or similar)	2·8	0·3	2·5

Thus, the significant differences between Groups M and O concern the mean numbers which were judged *correct* in the open-ended form of the test. In the cases both of identical and of similar wording, Group M offered significantly more correct responses than did Group O. As is to be expected, Group M also offered more ***incorrect*** multiple-choice phrases than did Group O.

Let us offer some illustrations, drawn from the stimulus word 'gullible.' The six phrases offered in the multiple-choice test were as follows: (i) island bird sanctuary, (ii) melting pot, (iii) easily cheated, (iv) can be swallowed without ill effects, (v) able to travel, (vi) gleeful. Of the twenty-seven members of Group M who scored one in the open-ended situation, seven gave 'easily cheated' (or 'easy to cheat'), six gave 'easily fooled,' the remaining fourteen giving such answers as 'taken in easily' or 'cheatable' or 'easily persuaded.' Twenty-three members of Group O scored one on this stimulus word, the majority of these being 'easily fooled' and 'easily taken in.' Other answers scoring one were 'believes anything,' 'naive' and 'green, one that always falls for practical joke.'

Examples of answers scoring half from members of Group O are: 'put one over falsely' and 'plyable' (*sic*). Answers scoring nought culled from Group O include 'truthful,' 'thick-skinned' and 'able to be eaten.'

The results presented so far indicate that in vocabulary testing the multiple-choice technique constitutes a substantially easier task for the subject than does the open-ended technique; and that taking the same vocabulary test, first in multiple-choice form and then in open-ended form, results in a significant number of multiple-choice answers previously seen appearing in the open-ended test. The latter tend to be correct answers, despite the fact that knowledge of results is not given.

The effect of taking the test in its multiple-choice form first can, however, be overestimated. Although the number of identical-plus-similar correct multiple-choice phrases given by Group M (8·4+11·2=19·6) exceeds the corresponding number given by Group O (2·2+4·0=6·2) by 13·4, it may be recalled that the mean number scoring one by Group M (17·4) exceeds the mean number scoring one by Group O (13·34) only to the extent of 4·06. Thus, about two-thirds of the number of M answers which are both correct and drawn from the multiple-choice test, are in fact, answered correctly by members of Group O but in their own, very different words.

C.—Relation between the subject's self-estimate and the correctness of his responses.

Let us now consider the relationship of the subject's own estimate of his word-knowledge to the correctness of his responses to the test items, (i) in the open-ended and (ii) in the multiple-choice situation. Table 4 shows (i) the percentage of words in the open-ended test, marked *a*, *b* and *c* which scored one, half, nought or were omitted, for Groups M and O, respectively; (ii) the percentages of words, in the multiple-choice test, marked *a*, *b* and *c* for which the correct solutions were chosen.

TABLE 4

THE RELATION BETWEEN THE SUBJECT'S SELF-ESTIMATE AND THE CORRECTNESS OF HIS RESPONSE:

(i) ON THE OPEN-ENDED TEST.

	Group M	Group O
Percentage of *a*'s scoring one in open-ended test	66·7	54·5
Percentage of *a*'s scoring half in open-ended test	13·4	20·3
Percentage of *a*'s scoring nought in open-ended test ...	15·9	20·3
Percentage of *a*'s omitted in open-ended test	4·0	4·9
Percentage of *b*'s scoring one in open-ended test.......	27·6	9·6 *sig.
Percentage of *b*'s scoring half in open-ended test ...	12·2	14·4
Percentage of *b*'s scoring nought in open-ended test ...	35·2	36·0
Percentage of *b*'s omitted in open-ended test	25·0	40·0 *sig.
Percentage of *c*'s scoring one in open-ended test	7·1	2·4 *sig.
Percentage of *c*'s scoring half in open-ended test	6·0	5·0
Percentage of *c*'s scoring nought in open-ended test ..	29·9	25·4
Percentage of *c*'s omitted in open-ended test	57·0	67·2 *sig

* The proportions in these cases differ significantly being between three and six times their standard error.

(ii) ON THE MULTIPLE-CHOICE TEST.

	Group M	Group O
Percentage of *a*'s correct	83·5	80·2
Percentage of *b*'s correct	57·8	52·7
Percentage of *c*'s correct	37·2	39·2

Three points of interest emerge from Table 4, (i) :

(*a*) The subjects do have a moderately good idea of their own word-knowledge. This is seen most clearly by inspection of those responses scoring one and of the ' omitteds.' On the other hand, (*b*) the ' halves ' and the ' noughts ' show a far less clear cut effect. Let us now consider (*c*) the results from the point of view of their Group (M or O). It looks as though taking the test first in its multiple-choice form (Group M) effects a striking improvement, for *a*'s, *b*'s and *c*'s, on the percentage of responses scoring one but has little effect on the percentage of responses scoring half or nought. For the omitteds, the differences for Groups M and O are significant for the *b*'s and *c*'s. It is possible that members of the M group gained confidence concerning their *b* and *c* words, by taking the test first in its multiple-choice form.

Part (ii) of Table 4 shows that, whilst the percentage of *a*'s correct is less than 100, that of correct *b*'s is slightly over the 50 per cent. which might have been predicted. The percentage of *c*'s correct is more than twice the one-in-six which would be expected by chance.

The differences between Groups M and O here are small. However, the differences between the multiple-choice percentages, on the one hand, and the relevant open-ended percentages, on the other, are seen to be very considerable.

D.—Association between intelligence test and vocabulary test.

Our last aim was to consider the relation between intelligence and vocabulary: to ascertain whether the correlation between AH 4 and the Self-judging Vocabulary test was influenced by the form in which the latter test was taken. Vocabulary tests with a steep gradient have repeatedly been found to correlate highly with intelligence tests. As is well known, many intelligence tests (such as the Binet and the Wechsler) incorporate a graded vocabulary test as one of the most favoured items.

The tests which constitute the Self-judging Vocabulary Scale, however, differ from the traditional vocabulary tests in that within each of its five separate tests, there is no gradient of difficulty. Therefore, whilst we predicted positive correlations between the S-j Test E and AH 4, we expected that they would be lower than is customary between vocabulary and intelligence tests. Indeed, we already possessed evidence to this effect (Heim and Watts, 1961).

TABLE 5

INTERCORRELATIONS OF THE THREE TESTS USED.

	Group M ($n=50$) r	Group O ($n=50$) r	Groups M+O ($n=100$) r
AH 4 with multiple-choice test ...	0·384 $p<0·01$	0·149 n.s.	0·265 $p<0·01$
AH 4 with open-ended test	0·438 $p<0·01$	0·187 n.s.	0·310 $p<0·01$
Multiple-choice with open-ended test	0·914 $p<0·01$	0·856 $p<0·01$	

The results shown in Table 5 are interesting in that AH 4 correlates significantly both with multiple-choice and open-ended vocabulary for Group M, whilst for Group O neither correlation is significant. Thus, it appears that the association between intelligence test score and vocabulary test may well be influenced by the form in which the first vocabulary test is taken and—

somewhat disconcertingly—that the association is decreased by presenting the vocabulary test first in the open-ended form. It is possible that the closer relationship with AH 4 for the M group is due to the fact that AH 4 and the M group's first vocabulary test share the multiple-choice technique. The (non-significant) increase of correlation between AH 4/open-ended over AH 4/multiple-choice may be due to the ability of the brighter subjects to benefit relatively more than the duller, from having taken the multiple-choice version of the vocabulary test immediately before the open-ended version.

The lowest line of Table 5, showing the correlations between multiple-choice and open-ended form for the two groups may be taken as measures of the consistency of the vocabulary test.

IV.—DISCUSSION AND CONCLUSIONS.

The four major questions posed in the Introduction have been dealt with. We have shown that: (*a*) The multiple-choice technique provides a significantly easier task then does open-ended answering, when groups matched on an intelligence test take the same vocabulary test in the two forms. Our set-up was similar to that of Raven (who evidently did *not* find this with his Mill Hill Vocabulary test*), in that we, too, impose no time-limit, we invite guessing and we offer a choice of six solutions, in the multiple-choice situation. On the other hand, the Self-judging Vocabulary tests have no built-in gradient of difficulty; and the six solutions offered are always substantially easier than are the stimulus words. This latter point may be particularly important—in this experiment and in vocabulary testing generally.

(*b*) The order in which two forms of vocabulary test are taken does have an effect on the subject's responses. One of the most interesting findings here is the tendency for members of Group M to give answers in the open-ended situation which are identical with, or similar to, the correct answer as given in the multiple-choice test, despite the fact that they were given no knowledge of results.

(*c*) The subjects' scores on both forms of the test proved to be fairly closely related to their self-estimates of word-familiarity, despite the fact that the groups tested were not very highly literate. Again, a big difference was found between the multiple-choice and the open-ended test results. For instance, 37 per cent. of the stimulus words marked *c* were answered correctly in the multiple-choice situation (Group M) whereas only 7+6=13 per cent. of the *c*'s (the ones plus the halves), gained positive scores for Group M in the open-ended situation. It is perhaps worth remarking that, although 13 per cent. is less than the one-in-six which would occur by chance in the multiple-choice, it is nearly twice as high as the 7 per cent. (2 ones plus 5 halves) obtained by Group O on their *c*'s in the open-ended test.

---

* Since writing this paper, the authors have received further data on the Mill Hill Vocabulary Test which suggests that Raven may be mistaken in his claim of a difference of one or two points, only, between multiple-choice mean and open-ended mean. With a group of 213 mature students, hoping to enter university or a college of education, the results were as follows:

	Mean	s.d.
Synonyms (multiple-choice) …………	31·8	3·42
Definitions (open ended) ………………	24·3	4·87

Difference = 7·5. This is 18 x standard error and, therefore, highly significant.

(*d*) The association between intelligence test and vocabulary test score has been shown to be considerably influenced, not so much by the form of vocabulary test as by the order in which these tests are taken, Group M yielding far higher correlations than did Group O.

Looking at the open-ended protocols, we gained the impression that our subjects, when unsure, were loth to guess—though they evidently had no objection to doing so in the less vulnerable situation of multiple-choice. Furthermore, when they did give wrong answers, these were sometimes due to the subject's confusing the stimulus word with another, often acoustically similar, word. For instance, in reply to LUSTROUS, one subject wrote 'famous' (presumably thinking of ' illustrious ') and another wrote ' to go forward in passion' (presumably thinking of ' lustful '). The sailor who offered ' film equipment ' in reply to SINECURE was, presumably, unworried about the niceties of s's and c's ; and there is little doubt about the cause of confusion in the following answer to POSTERITY : 'The standard of living in Great Britain in the last ten years has gone up—this is posterity.'

On the other hand, there were some wholly incomprehensible wrong answers which had a certain appealing plausibility, such as OPAQUE : ' the measurement in liquids,' and PECUNIARY : ' a part of the heating in the body.' These may, of course, have been leg-pulls or they may be the results of forcing oneself to guess in the absence of any clues.

ACKNOWLEDGEMENTS.—We wish to express our thanks to Mr. E. G. Chambers for statistical assistance ; to Mrs. V. Simmonds for doing most of the testing ; and to the Medical Research Council for financial support.

V.—REFERENCES.

COOK, J. M., HEIM, A. W., and WATTS, K. P. (1963). The word-in-context : a new type of verbal reasoning test. *Brit. J. Psychol.*, 54, 227-237.

HEIM, A. W. (1955). *Manual for the Group Test of General Intelligence AH* 4. London : National Foundation for Educational Research.

HEIM, A. W. and WATTS, K. P. (1961). A preliminary study of the Self-judging Vocabulary Scale. *Brit. J. Psychol.*, 52, 175-186.

RAVEN, J. C. (1958). *Guide to using the Mill Hill Vocabulary Scale.* London . H. K. Lewis and Co., Ltd.

(*Manuscript received 23rd August,* 1966)

TEACHERS' RATINGS OF SCHOOLBOYS' BEHAVIOUR RELATED TO PATTERNS OF SCORES ON THE NEW JUNIOR MAUDSLEY INVENTORY

BY H. B. GIBSON

(*Research Psychologist, Institute of Criminology, University of Cambridge*)

SUMMARY. The New Junior Maudsley Inventory was administered to 394 schoolboys, of late primary school age, for whom teachers' behavioural ratings were also available. A series of comparisons were made between the two groups of boys rated as particularly well behaved, and particularly badly behaved. While the standard scales of the inventory had some power to discriminate between the groups, item analysis of the data gave new scales which were more meaningful for this population and which could be combined to give superior powers of discrimination. There was some evidence that the trait of extraversion could be considered as having a dual nature, and that the Lie scale represented not only a defensive 'priggishness,' but an inventory of an objective nature for self-reported delinquency.

I.—INTRODUCTION.

TEACHERS' ratings of behaviour in school are about the most reliable measure we have of children's general social adjustment, and such ratings are excellent predictors of future delinquency and neurotic maladjustment (e.g., Mulligan, *et al.*, 1963). Some early work of Burt, which has been recently reported (Burt, 1965), indicates that teachers' ratings are related both to personality factors of extraversion and neuroticism and also to delinquent behaviour many years later.

The theories of Eysenck regarding personality dimensions in relation to delinquent behaviour and social maladjustment hardly need re-stating; however, it cannot be said that these theories have yet been adequately tested. A long series of empirical studies with adult and adolescent offenders (e.g., Bartholomew, 1959; Field, 1959; Fitch, 1962; Little, 1963) have failed to arrive at any satisfactory conclusion.

Apart from Eysenck's own pertinent comment on the last-mentioned study that it, and similar studies, are hardly a test of his theories if the subjects are highly selected samples of a selected population (Eysenck, 1963) there are methodological objections to the use of the results of personality inventories in the manner followed in these studies. One gets the impression that these research workers had not grasped a rather mundane fact—that a questionnaire item may convey very different meanings to different people, particularly if they are answering under rather special conditions. Thus, if the Maudsley Personality Inventory (M.P.I.) forms which were completed by the inmates of prisons and borstals had been subjected to factorial analysis, while it is almost certain that scales of extraversion, neuroticism and lying would have emerged, it is by no means certain that the items associated with those scales would have been the same as in the standard scales of the M.P.I. To take a concrete example, affirmation of the M.P.I. item " Do you *always* answer a personal letter as soon as you can after you have read it ?" means an increment to the L score, and rightly so for most normal populations. However, if immediate reply to letters is objectively true for most prisoners, as it may well be under prison conditions, then it is absurd to let affirmation increment the L scale. Reading the inventory one may suspect on *a priori* grounds, that the meaning of many items will be quite different for a prison population ; the question of whether this suspicion is justified or not *could* have been settled by proper item analysis of the results.

Ideally, we want tests with scoring keys and norms which are applicable in all circumstances. This is probably an impossible ideal as far as personality tests are concerned. In these days of easy computer facilities, however, we have the remedy, the factorial analysis of the results obtained from any large, special population. This would reveal just what clusters of items went to make up recognizable scales, cut down the error variance and ensure at least the construct validity of such terms as extraversion and neuroticism applied to the particular population under scrutiny.

II.—The present study.

This study was not designed to validate (or invalidate !) Eysenck's theories. It was merely part of a process of gaining as much useful information as possible about a population of schoolboys who are the subject of a follow-up study. The implications for fundamental research in personality theories are of importance, no doubt, but the study has been conducted on pragmatic lines.

III.—Method.

The New Junior Maudsley Inventory (N.J.M.I.) was administered to 394 boys in schools. With a few exceptions, the boys were aged 10 to 11 years and in the last term at their primary schools. The administration was in classroom groups, with two or three psychologists invigilating, and the teacher out of the room. Because this population contained a number of rather poor readers, after the inventory had been explained each item was read out aloud, and it was pointed out that complete frankness would do no harm as the forms were not going to be shown to anyone connected with the schools. In a minority of cases, the inventory had to be administered individually because certain boys were absent on the day of group testing.

There were available for all these children ratings on their behaviour over the year, made by the school teachers on a form that has been described elsewhere (Gibson, 1964a).

IV.—Results.

On the basis of the teachers' ratings, the boys were divided into three groups. A (' good '), B (' average ') and C (' naughty ') in the standard manner (Gibson, *op. cit.*).

The inventories were scored according to the published norms of the test, and the relationship of the E and the N scales to the status of A and C boys is shown in Table 1. On these 2 × 2 tables, taking the cutting points as shown, the C boys are relatively more extraverted and more neurotic than the A boys. However, the N scale appears to be a far more powerful discriminator than the E scale, and considering the general psychological meaning given to these two personality variables, this is not necessarily an absolute finding.

Further analysis.

All the sixty-two items of the Inventory were intercorrelated and analysed by the method of Principal Components. Four components were extracted, accounting for 8·0 per cent., 5·3 per cent., 3·7 per cent. and 3·1 per cent. of the variance.

Component 1 was clearly asscoiated with the L scale. The L scale items had loadings averaging ·43 on it, and all were over ·27, with the exception of item 44 (" I am a bit afraid of policemen ") which had a loading of only ·14 on it, an indication that it could reasonably be dropped from the scale for this population. Component 2 was associated with the N scale. Components 3 and 4 were associated with the E scale, but whereas the former was associated with social extraversion (e.g., "I'd rather go to a party than stay at home ") the latter appeared to be associated more with items indicative of boldness *versus* lack of self-confidence (e.g., " I hate to have people look at me when I'm working ").

A graphical plot of the loadings of all items on the axes of the first two components was prepared, and the axis of Component 2, was rotated to maximize the loadings of the N scale items. This necessitated rotation through an angle of 15°. The item with the highest loading (·58) on the resultant C'_2 axis was, " I am always afraid that sad things will happen to me," which may be taken as an indicator of the sort of neuroticism which is implied by this vector. A further graphical plot was made showing the loadings of the items on C'_2 and C_4. From this it was possible to pick out twenty-one items for the extraversion scale and sixteen items for the neuroticism scale. These two shortened scales will henceforward be referred to as E' and N'. Comparing the old scales with the new scales developed *ad hoc* from the results of testing this population, two items of the old E scale were now incorporated in the N' scale, and five items of the N scale were now incorporated in the E' scale.

Table 2 shows the relation of the new scales of E' and N' to the extreme categories of the teachers' ratings. Among the stable introverts A boys predominate 8 : (1), and among the neurotic extraverts C boys predominate 3 : (9). In the cell representing stable extraverts the figures are equal, 8 : (8) ; however, a study of the raw data revealed that all boys with E' scores over 14 were in fact C boys. As the scale of E' appears to have about twice the discriminative power than N' does, an actuarial sort was carried out giving each boy a score according to the formula of 2E'+N'. The results of this sort are shown in Table 3.

The Lie Scale.

The new lie scale L' was obtained by dropping one item, as already explained. This scale correlated only slightly with the other two scales ($r = -·11$ and $-·07$ with E' and N', respectively). This was because all scales were derived from orthogonal axes obtained by analysing the data for this population. In previous studies L scales have been shown to give quite substantial negative correlations with measures of N: e.g., Gibson (1962; 1964c), Eysenck; S. B. G. (1965). However, examination of the raw data suggested that there was, in fact, an interesting non-linear interaction effect between the L' scale and the N' scale. This is shown in Figure 1. It may be seen that whereas for subjects with N' scores of 5 and over ($n=157$) there is a unimodal distribution on the L'scale, for those subjects with lower N' scores ($n=137$) the distribution is distinctly bi-modal.

Among those scoring specially low on the N' scale, there is a recognizable group who score very high on the L' scale. This tendency is especially marked among the sub-population of ' good ' boys, as shown in Figure 2.

The relationship of the L' scale to the teachers' behavioural ratings is shown in Table 4. In the lowest array of L scores, the percentage of ' naughty ' boys is over double that of ' good ' boys, and at the other end of the L' scale the position is reversed. The difference is significant at the ·05 level of probability.

It is of interest to note that certain individual L scale items have a very high power of discrimination between A and C boys. The relative power of discrimination of the L scale items is shown in Table 3. Out of the nine items worded in the positive direction of social desirability or ' priggishness, ' seven have been endorsed by a higher percentage of A than C boys.

Final discrimination.

As all three scales have been shown to have a power to discriminate between A and C boys, a final joint discrimination on an actuarial basis was attempted. The scale of (2E'+N') was added to by a points score from the L' scale : viz., 0 to 3=2, 4 to 12=1, 13 to 17=0. The final discrimination on the basis of this score is shown in Table 6. This may be contrasted with Table 3 based on E' and N' alone. It may be seen that the addition of the L' scores raises the value of χ^2 from 13·50 to 16·61 and hence provides a superior level of discrimination.

V.—Discussion.

This research has shown that a standard personality inventory applied to a fairly large and homogeneous population of boys, could be improved in its discriminative power by a process of reconstruction in terms of item selection and grouping using the actual data obtained. According to the published norms, neuroticism was significantly associated with bad behaviour in the boys of this population, but extraversion was only slightly associated. The factorial analysis of the data revealed that extraversion could be considered in terms of two separate components, social extraversion-introversion and boldness-shyness. This question has already been discussed by Eysenck and Eysenck (1963). In terms of the relationship of extraversion to bad behaviour in school boys, it is unquestionably the aspect of boldness and defiance rather than that of social extraversion which must be taken into consideration.

The type of boy who was most typical of those whom the teachers rated as ' naughty ' gave a pattern of scores which can rightly be labelled ' neurotic extraversion.' However, it is of interest to note that at the very extreme of the E' scale the N' scores were low, but that the chances of the extremely extraverted boys being rated as ' naughty ' was very high.

In the present study, the population tested was not so very different in character from that on which the scales were originally standardized, but as the administration both to the standardization population and to this experimental population was personally conducted by the author himself, he is only too conscious of some important differences between the populations. Perhaps the most important difference was that of social class and level of culture. A test must have norms, and that those of the N.J.M.I. are fairly robust, has been shown in many studies, but there is no substitute for investigating one's data within the context in which it is collected. The amendments to the scales in this study, will be of no real help to any other study. Their relevance is only to this particular population tested at this particular time and place.

This study showed that the L scale was the most important of the three, since it accounted for the largest portion of the variance. Its meaning is still unclear. The results have contradicted a suggestion arising out of a pilot group (Gibson, 1964b) that badly behaved boys would, on the whole, tend to give high L scores. In this population, at least, a high degree of ' priggishness ' seems to be most characteristic of the ' good ' boys. A *very* low score does, however, as suggested before, seems to indicate a defiant boasting about naughtiness, and to be most characteristic of the badly behaved boys. The bimodal distribution of boys who were particularly low on N' (especially among the

' good ' boys) probably indicates that the lower mode is a fair approximation to the truth about their moral conduct and that the upper mode indicates a tendency to lean over backwards, in presenting a good picture of themselves.

The L scale may, in fact, serve two separate purposes. First, as an indicator of the degree to which the boy is 'lying ' about himself, and a *very* low L score may be as indicative of attempted deception as a very high score. Second, as an inventory of self-reported delinquency. Table 5 shows that the badly behaved boys admitted to naughty acts more frequently, and in certain cases the level of discrimination was strikingly high (e.g., item 21, which concerns stealing) and was certainly in accordance with objective reality.

VI.—REFERENCES.

BARTHOLOMEW, A. A. (1959). Extraversion-introversion and neuroticism in first offender and recidivists. *Brit. J. Delin.*, 10, 120-129.

BURT, C. (1965). Factorial studies of personality and their bearing on the work of the teacher. *Brit. J. Educ. Psychol.*, 35, 368-378.

EYSENCK, H. J. (1963). Comment on Dr. Little's paper. *Brit. J. Crim.*, 4, 160-162.

EYSENCK, S. B. G. (1965). A new scale for personality measurements in children. *Brit. J. Educ. Psychol.*, 35, 362-367.

EYSENCK, S. B. G., and EYSENCK, H. J. (1963). On the dual nature of extraversion. *Brit. J. Soc. Clin. Psychol.*, 2, 46-55.

FIELD, J. G. (1959). The personalities of criminals. *Bull. Brit. Psychol., Soc.*, 38, 22a.

FITCH, J. H. (1962). Two personality variables and their distribution in a criminal population : an empirical study. *Brit. J. Soc. Clin. Psychol.*, 1, 161-167.

GIBSON, H. B. (1962). The lie scale of the Maudsley Personality Inventory. *Acta Psychologica*, 20, 18-23.

GIBSON, H. B. (1964a). The Spiral Maze : a psychomotor test with implications for the study of delinquency. *Brit. J. Psychol.*, 54, 219-255.

GIBSON, H. B. (1964b). The validity of the lie scale of a children's personality inventory. *Acta Psychologica*, 22, 241-249.

GIBSON, H. B. (1964c). A lie scale for the Junior Maudsley Personality Inventory. *Brit. J. Educ. Psychol.*, 34, 120-124.

LITTLE, A. (1963). Professor Eysenck's theory of crime : an empirical test on adolescent offenders. *Brit. J. Crim.*, 4, 152-160.

MULLIGAN, G., DOUGLAS, J. W. B., HAMMOND, W. A., and TIZARD, J. (1963). Delinquency and symptoms of maladjustment : the findings of a longitudinal survey *Proc. Royal Soc. Med.*, 56, 1083-1086.

(*Manuscript received* 12*th October*, 1966)

TABLE 1

DISCRIMINATION BETWEEN A ('GOOD') and C ('NAUGHTY') BOYS BY THE E AND N SCALES, AS PUBLISHED.

	Extraversion	
	<13	>13
A boys	51	36
C boys	45	48

$\chi^2 = 1\cdot50$ N.S.

	Neuroticism	
	<10	>10
A boys	62	25
C boys	43	50

$\chi^2 = 10\cdot58$ with 1 df. $p < \cdot01$

TABLE 2

DISCRIMINATION BETWEEN A BOYS AND C BOYS BY THE E' AND N' SCALES DERIVED IN THIS STUDY.

The figures in each cell show the frequencies. Figures in parentheses indicate the C boys.

	N' Score		
E' Score	—3	4—8	9+
—7	8 (1)	16 (9)	3 (1)
8—11	12 (4)	28 (29)	4 (11)
12+	8 (8)	5 (21)	3 (9)

$\chi^2 = 27\cdot8$ with 8 df. $p < \cdot001$.

TABLE 3

DISCRIMINATION BETWEEN A BOYS AND C BOYS BY THE SCALE DERIVED FROM 2E'+N'.

	Score Arrays							
	—15	—18	—21	—24	—27	—30	—33	34+
A boys	8	11	15	15	17	13	6	2
C boys......	2	5	7	17	14	20	12	16

By the Kolmogorov-Smirnov test, $\chi^2 = 13\cdot5$ with 2 df. $p \simeq \cdot001$.

TABLE 4

Discrimination between A boys and C boys by the L' Scale.

	L' Scale						
	0—3	4—5	6—7	8—9	10—12	13+	Total
A boys	Percentage 7 (8·1)	Percentage 16 (18·4)	Percentage 17 (19·5)	Percentage 11 (12·6)	Percentage 18 (20·7)	Percentage 18 (20·7)	Percentage 87 (100)
C boys	16 (17·2)	12 (12·9)	29 (31·2)	16 (17·2)	12 (12·9)	8 (8·6)	93 (100)

On the Kolmogorov-Smirnov test $\chi^2=7.12$ with 2 df. $p<\cdot 05$.

TABLE 5

Percentage of Boys Endorsing ' Same ' Against the Various Items of the L' Scale.

Numbering of the items as in the published N.J.M.I.

Item No.	A boys (' good ')	C boys (' naughty ')	Difference
7	57·5	46·7	10·8
10*	63·4	55·4	9·0
14*	56·3	30·0	26·3
17	39·1	47·8	—8·7
21*	67·8	46·7	21·1
24*	42·5	37·0	5·5
28	56·3	78·3	—22·0
32	52·9	63·0	—10·1
35*	44·8	49·0	—4·2
38	59·8	70·7	—10·9
41*	50·6	48·9	1·7
47*	59·8	55·4	4·4
50	58·6	71·7	—13·1
53*	62·1	64·1	—2·0
56	64·4	69·6	—5·2
59*	47·1	44·6	2·5
63	55·2	62·0	—6·8

* Items which are worded in the ' priggish ' direction.

TABLE 6

FREQUENCY DISTRIBUTIONS OF A AND C BOYS ON THE JOINT SCORE DERIVED FROM E', N' AND L'.

	Score Arrays								
	–15	–18	–21	–24	–27	–30	–33	–36	37+
A boys	7	5	17	14	19	15	5	3	2
C boys	2	2	6	10	22	12	19	11	9

On the Kolmorgorov-Smirnov test $\chi^2 = 16\ 6$ with 2 df. $p < \cdot 001$.

FIGURE 1

FREQUENCY DISTRIBUTIONS ON THE LIE SCALE OF TWO GROUPS DISTINGUISHED BY THEIR LEVEL OF NEUROTICISM SCORE. TOTAL POPULATION.

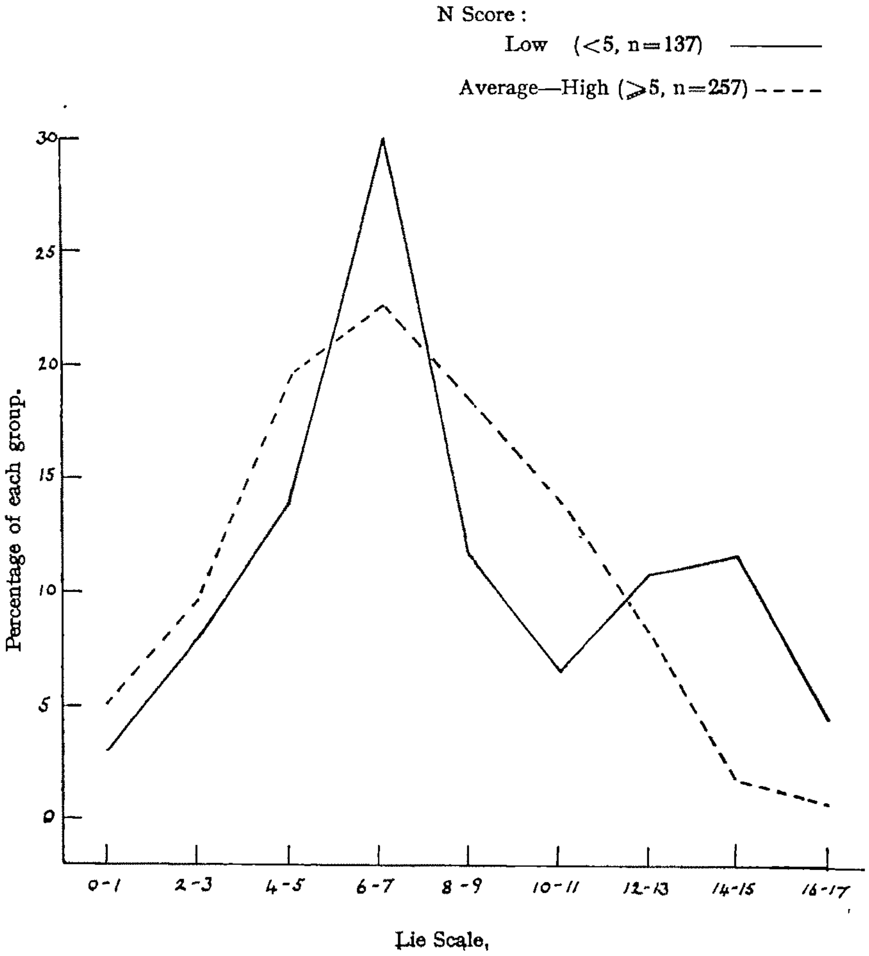
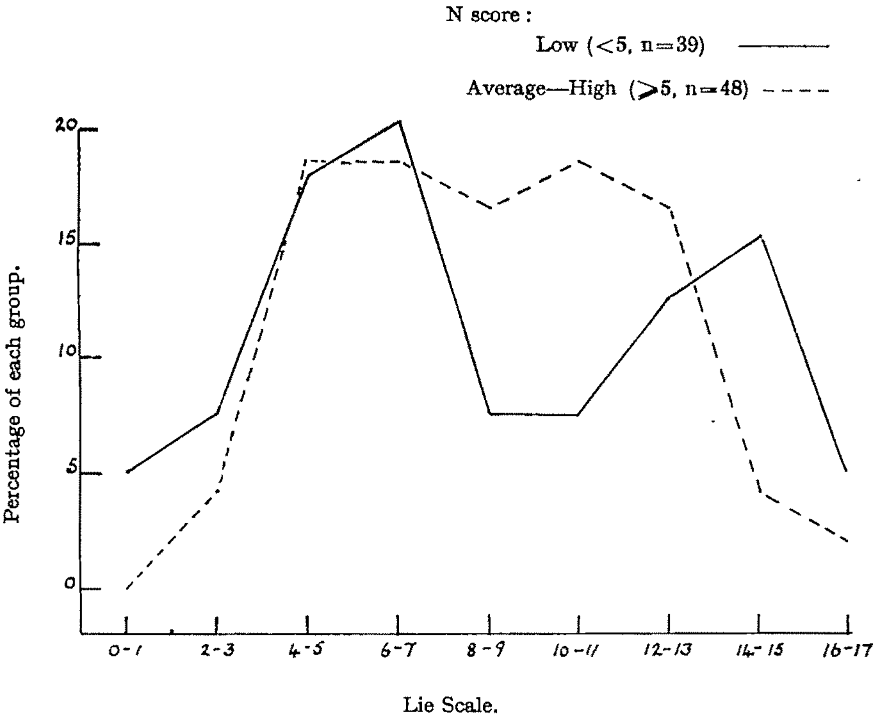


FIGURE 2

FREQUENCY DISTRIBUTIONS ON THE LIE SCALE OF TWO GROUPS DISTINGUISHED BY THEIR LEVEL OF NEUROTICISM SCORE. "GOOD" BOYS ONLY.



PRODUCTIVE THINKING IN RETARDED AND NON-RETARDED CHILDREN*

BY JOHN F. CAWLEY AND DONNA V. CHASE
(*University of Connecticut*)

SUMMARY. Three categories of productive thinking measures were administered to retarded and non-retarded children of equal mental age. There were no significant differences between retarded children in special class, retarded children in regular class and average children in regular class. The results are discussed in terms of future research and the implication for teacher-training programmes.

I.—INTRODUCTION.

THROUGHOUT a considerable portion of our present century, mentally retarded children have been perceived as poor learners, concrete thinkers, and lacking in abstract or creative ability. Within the past decade, however, research has been conducted in which the performance of mentally retarded, average and above average children has been compared on a variety of tasks. A reasonable, although limited conclusion appears to be that if children of different intellectual levels are contrasted when selected on a basis of mental age, greater similarity exists in their performance than was previously acknowledged. This should not infer that differences in performance will be eliminated if the developmental status between groups is equated ; it does infer that the discrepancies are not as great and that our generalizations pertinent to the mentally retarded may be in need of further clarification.

A perusal of selected curiculum guides for teachers of retarded and non-retarded children indicates that concrete and limited experiences are commonly advocated for use with the retarded, while non-mentally-retarded children are encouraged to explore many avenues to learning and problem solving. It appears that mentally retarded children are trained to be convergent thinkers, whereas normal children are urged to demonstrate greater use of flexibility and originality in their thinking patterns.

Recent research (Tisdall, 1962 ; Rouse, 1965 ; Taba, 1964) has focused attention upon the productive thinking characteristics of the mentally handicapped ; the present investigation differs from it in that the children selected did not benefit from any specific programme designed to improve their productive thinking qualities, as was characteristic of the others. Furthermore, the present project identified three categories of stimulus/response modes which were utilized in the assessment of productive thinking traits. These were : (1) Verbal Stimuli requiring Verbal Responses ; (2) Non-Verbal Stimuli requiring Verbal Responses ; (3) Non-Verbal Stimuli requiring Non-Verbal Responses.

The purpose of the present study was to compare retarded and non-retarded children of equal mental ages on a variety of productive thinking tasks.

---

* The research reported herein was conducted pursuant with a contract with the Division of Elementary and Secondary Research, Office of Education, Project No. 5-8106-2 12-1.

II.—LITERATURE.

There exists a fair amount of literature which has a generalized relationship to the topics under discussion, much of it stemming from the efforts of Guilford (1950). The specific literature is not so nearly abundant. Cawley and Chase (1966) failed to identify any pattern of significant correlations between M.A., I.Q. and the productive thinking characteristics of retarded and non-retarded children. Torrance (1962), upon whose procedures the present study was developed discusses at length, the characteristics of creative teachers and he suggests modifications in the teacher-education programme which might bring about maximum creativity in teachers. The role of the teacher appears important and there is research in the area of mental retardation which strongly suggests that instructional practices can be developed which will positively influence the productive thinking abilities of mentally retarded children. Tisdall (1962) compared a group of retarded children enrolled in regular classes, a group of mentally retarded children in special classes and a group of average children on a variety of productive thinking tasks. The results showed that special class retarded children and average children in regular classes performed significantly higher than retarded pupils in regular classes on verbal tasks ; no differences were found on non-verbal tasks. The experimental teachers of the mentally retarded children in special classes consistently employed a discovery method of teaching, and encouraged their students to formulate ideas, dramatize these ideas and participate actively in the presentation of new thoughts and notions.

Additional data have been obtained from an experiment which studied the effects of a training programme on the productive thinking of mentally handicapped children (Rouse, 1965). Experimental subjects were presented with a training sequence which included thirty lessons of approximately thirty minutes each. Assessment of the effectiveness of the programme was undertaken via pre-and post-test comparisons of experimental and control groups on the Product Improvement and Circles Tasks of the *Minnesota Tests of Creative Thinking*. The mean gain for the experimental groups was substantially greater than that for controls.

Trained teachers were able to raise the level of thought processes used by students in their classes, even to the point that some children of low I.Q., performed on the same level of abstraction as children with high I.Qs. (Taba, 1964, cited by Amidon and Sewin, 1965).

The literature relevant to the present study shows that (1) productive thinking can be nurtured in retarded and average children ; (2) instructional practices tend to influence the expression of productive thinking in children ; and (3) children of low I.Q. can perform toward an abstract level of thinking.

III.—PROCEDURE.

A total of seventy-eight children participated in this project. This population was composed of a sample of mentally retarded children enrolled in *special classes* in public schools, a sample of retarded children enrolled in regular classes, and a sample of non-retarded children enrolled in regular classes in public schools. The characteristics of the three samples are presented in Table 1. Although differences in intelligence quotients are noted, Table 2, the major criterion used for selections, mental age, is shown to be a non-significantly differentiating characteristic among the three groups, as determined by the *California Test of Mental Maturity*.

TABLE 1

CHARACTERISTICS OF SUBJECTS IN PRESENT STUDY: INTELLIGENCE QUOTIENTS AND MENTAL AGES.

		Special Class Retarded (N=26)	Regular Class Retarded (N=26)	Regular Class Non-Retarded (N=26)
Intelligence Quotient	$\bar{X}$	64·92	71·73	93·50
	Range	60–77	58–80	89–102
Mental Age in Months	$\bar{X}$	111·69	114·96	110·19
	Range	103–124	92–125	103–119

TABLE 2

COMPARISON OF CLASS GROUP MEANS ON INTELLIGENCE QUOTIENTS AND MENTAL AGES.

	Intelligence Quotient "t"	Significance Level	Mental Age "t"	Significance Level
SCR—RCR	·84	N.S.	·28	N.S.
SCR—NRRC	4·17*	$P<\cdot01$	·15	N.S.
RCR—NRRC	2·86*	$P<\cdot01$	·46	N.S.

* Statistically Significant.

Special Class Retarded=SCR.
Regular Class Retarded=RCR
Non-Retarded Regular Class=NRRC.

On an individual basis, a battery of productive thinking tests was administered to each subject. These tests, categorized according to stimulus/response mode, are listed as follows:

(1) *Verbal Stimuli Requiring Verbal Responses*:
  (*a*) Unusual Uses Problems (Tin Cans).
  (*b*) Improvement Task.
  (*c*) Mother Hubbard Problem.

(2) *Non-Verbal Stimuli Requiring Verbal Responses*:
  (*a*) The-Ask-And-Guess-Test.
  (*b*) Product Improvement Task (Toy Dog).
  (*c*) Unusual Uses (Monkey).

(3) *Non-Verbal Stimuli Requiring Non-Verbal Responses*:
  (*a*) Incomplete Figures Task.
  (*b*) Circles Test.
  (*c*) Picture Construction Task (Triangle).

Verbal responses were recorded by the examiner and tabulated. Responses to non-verbal tasks were also acquired and tabulated. All scoring was accomplished in accordance with the procedures advocated by the developers of the test materials.

IV.—Results.

The means, standard deviations and F ratios by stimulus/response modes are presented in Table 3. An examination of the data indicate that there are no overall significant differences among the groups on any of the three test modes. Mentally retarded children in special classes, mentally retarded children in regular classes and non-retarded children of equal mental ages do not appear to be significantly different in their productive thinking abilities under conditions of Verbal Stimuli—Verbal Response; Non-Verbal Stimuli—Non-Verbal Response ; Non-Verbal Stimuli-Verbal Response.

TABLE 3

Means, Standard Deviations and F Ratios for Three Treatment Groups on Three Stimulus/Response Mode Combinations of a Test of Productive Thinking.

Stimulus/Response Mode		Special Class Retarded	Regular Class Retarded	Average Class Non-Retarded
Non-Verbal S	$\bar{X}$	65·08	64·46	65·85
Non-Verbal R	S.D.	28·09	31·14	16·70
	N	26	26	26
	*F= $\frac{12{\cdot}52}{653{\cdot}28}$ = ·02 · NS			
Verbal S	$\bar{X}$	29·42	31·38	35·69
Verbal R	S D.	17·35	22·26	24·40
	N	26	26	26
	*F= $\frac{217{\cdot}39}{446{\cdot}13}$ = ·49 . NS			
Non-Verbal S	$\bar{X}$	22·73	26·92	30·50
Verbal R	S.D.	11·04	18·47	17·26
	N	26	26	26
	*F= $\frac{393{\cdot}16}{243{\cdot}86}$ =1·61 : NS			

* Based on one-way analysis of variance with 2 and 75 degrees of freedom.

V.—Discussion.

In contrast to other research relative to productive thinking abilities among the mentally retarded (Tisdall, 1962 ; Rouse, 1965), this project was undertaken with subjects who had not benefited from an instructional programme specifically aimed at improving these traits. There were no significant differences among the three treatment groups on any of the three stimulus/response modes. However, without adequate normative data it is impossible to establish a reference point from which the performance of the subjects in this project can be contrasted. Accordingly, the overall performance levels identified cannot be described on a qualitative basis. It is assumed, therefore, that the performance of the subjects in the present study is representative of the population from which they were selected and the suggestion is tendered that curriculum plans for the mentally retarded should not exclude productive thinking activities.

Assuming that the retarded youngster is capable of demonstrating a reasonable level of productive thinking abilities, it becomes essential for us to determine the relationship of these qualities to such factors as language development, academic performance, and social or vocational success. Is it possible that certain types of productive thinking are more susceptible to improvement than others ? Does this result in a generalized change in the adaptive behaviour of the retarded ? Are there certain types of productive thinking patterns which are more identifiable among retarded children of one age, ability or developmental level than there are at others ? What predictive validity do measures of productive thinking attain ? The aforementioned are but a few of the many issues which should be considered within the field of mental retardation.

Teachers of the mentally retarded should be exposed to the notion of productive thinking in their training programmes. Planning for instruction ought to include the incorporation of productive thinking activities into the classroom. Teachers should have sufficient contact with the processes to enable them to stimulate the children toward a particular type of productive thinking and to assess the response patterns of the pupils. In effect, teaching thinking requires planning and the literature cited herein, and the results of the present study, suggest that the mentally handicapped are entitled to such planning.

VI.—REFERENCES.

CAWLEY, J. F., and CHASE, D. V. (1966). *Productive Thinking in Retarded and Non-Retarded Children.* Washington, D.C.: U.S. Department of Health, Education and Welfare : Office of Education, Project No. 5-8106-2-12-1.

CLARK, D., VELDMAN, E., and THORPE, J. (1965). Convergent and divergent thinking abilities of talented adolescents. *J. Educ. Psychol*, Vol. 56, pp. 157-163.

GETZEL, J. W., and JACKSON, P. W. (1960). *The Study of Giftedness : A Multidimensional Approach.* The Gifted Student. Washington, D.C.: U.S. Department of Health, Education and Welfare, U.S. Government Printing Office.

GUILFORD, J. P. (1950). Creativity. *Amer. Psychol*, Vol. 5, pp. 444-454.

ROUSE, SUE (1965). Effects of a training program on the productive thinking of educable mentally retardates. *Amer. J. Mental Deficiency*, Vol. 69, pp. 666-673.

TABA, H., and others (1964). *Thinking in Elementary School Children.* San Francisco, California : U.S. Department of Health, Education and Welfare ; Office of Education, Coop. Research Project No. 1574.

TISDALL, WILLIAM (1962). Productive thinking in retarded children. *Exceptional Children*, Vol. 29, pp. 36-41.

TORRANCE, E. P. (1962). *Guiding Creative Talent.* New Jersey : Prentice-Hall, Inc., Englewood Cliffs.

(*Manuscript received 3rd November*, 1966)

ABILITY IN SCIENCE AT ORDINARY LEVEL OF THE GENERAL CERTIFICATE OF EDUCATION

BY D. G. LEWIS

(Department of Education, University of Manchester)

SUMMARY. Objective tests of ability in physics, chemistry and biology were constructed, each test being structured into the three broad categories of (i) basic knowledge, (ii) comprehension and application, and (iii) evaluation, along the lines of the Bloom Taxonomy. The tests were administered together with tests of verbal, numerical and spatial ability to all the boys who had completed the Ordinary Level science syllabuses of the Joint Matriculation Board in five schools in the Manchester area. A factorial analysis yielded the following results:

(1) A factor of scholastic achievement was obtained. This factor was of major importance in the categories of knowledge and of comprehension and application in each of physics, chemistry and biology.

(2) The categories of evaluation were linked factorally to but a limited extent to the categories of knowledge and comprehension and application.

(3) An aptitude or reasoning factor, comprising ability in the spatial, numerical and (to a less extent) verbal fields, was indicated. This factor contributed substantially to the categories of evaluation in physics, chemistry and biology.

(4) Numerical and verbal factors were also obtained. These factors, however, made little contribution to the measures of ability in science.

Some implications for the teaching of science in schools are discussed.

I.—INTRODUCTION.

FACTORIAL studies of attainment in school science (Ellis, 1948; Wrigley, 1956; Lewis, 1961) indicate that science would appear as a minor group factor which, together with a similar factor for mathematics, contribute to a major group factor opposed to one linking together most of the arts subjects. Factorial studies of science with psychological tests (Pawley, 1937; Ormiston, 1939; Pollard, 1942 Berridge, 1947; James, 1951; Lamb, 1954; Jog, 1955) suggest that ability may be influenced not only by the general cognitive factor (*g*) and a general personality factor (such as Alexander's *X*) but also by certain cognitive group factors representing verbal ability (*v*), spatial ability (*k*) and numerical ability (*n*). Such studies have for the most part, however, been based on measures of scientific ability modelled on the traditional school examinations. This is unsatisfactory in so far as undue stress has been laid on rote learning to the detriment of other, more essential, aspects of scientific ability. While a report of the Science Masters' Association, as far back as 1938, advised that both the teaching and examining of science should be concerned with scientific method no less than with the acquisition of facts, nevertheless, the implementation of this advice—at any rate as far as the formulation of examination papers is concerned—has been lacking (see, for example, the Gulbenkian Foundation Report, 1959; Crossland and Amos, 1961, 1965). A more comprehensive coverage of ability in science is clearly desirable.

II.—CATEGORIES OF SCIENTIFIC ABILITY.

A useful starting point is the Science Masters' Association Report (1938) to which reference has been made. The two major categories in the S.M.A. list are the *acquisition of scientific information and knowledge* (including knowledge of empirical facts, technical terms, laws and principles) and the *development of scientific modes of thought*—this second category, including the ability to distinguish facts from hypotheses, to plan experiments in order to test hypotheses and to draw inferences from data. A similar statement of these ' higher level ' aspects of scientific ability was made by Frutchley and Tyler (1937), whose list of ' desirable outcomes ' for science courses closely parallel that of the S.M.A.

Concurrent to the formulation of objectives was the construction of tests to measure at least some of the non-factual elements. Thus, the test of Engelhart and Lewis (1941) may be noted as an attempt to assess the ability to draw reasonable generalizations from experimental data. Again, Teichman (1944) formulated tests specifically to measure (i) the ability to draw a conclusion from given facts, (ii) the ability to select the best conclusion from alternatives, and (iii) the ability to evaluate proposed conclusions (and in particular, to decide whether they are reasonable, sufficient or pertinent with respect to the data).

Both the previous studies emphasize aspects of scientific ability other than the acquisition of information, and which may be placed under the broad heading of scientific method. A comprehensive survey of the many and varied aspects of scientific method was made by Keeslar (1945), and from the opinions of the practising scientists consulted a list of the essential elements was produced. This list, which includes such categories as the inventing and planning of experiments to test an hypothesis, the carrying out of the experiment and running check experiments involving the same experimental facts, was compiled with the needs of university students in mind, and might seem unduly ambitious as a programme for science courses in schools. It does nevertheless, provide a broad basis for varied teaching approaches, and for the construction of tests. Again, some of the categories are substantially the same as those in the previous studies.

Heil, Kambly, Mainardi and Weisman (1946) stressed the development of understanding in the context of the material and methods of science. Two broad categories of understanding are distinguished, that of *factual understanding*, covering such aspects as giving specific examples of scientific concepts, and using facts to make predictions and explain phenomena, and that of *method understanding*, involving such aspects as making proper qualifications when interpreting data, identifying unstated assumptions and valid cause-and-effect relationships. No distinction as to level (i.e. high school or university) is made, though from the illustrative test items it would seem that, in the opinion of the writers, both categories of understanding could be taught, and tested, in schools.

In a study of the abilities involved in scientific thinking, Burmester (1952) produced a list which parallels to a considerable extent Keeslar's list of the elements of scientific method, as well as covering most of the two broad categories of the study by Heil, Kambly, Mainardi and Weisman. Thus, Burmester cites the ability to recognize a hypothesis, to plan experiments to test a hypothesis, to interpret data (e.g., to evaluate the relevance of data, and to distinguish between facts and inference) and to apply generalizations to new situations. In a subsequent paper (1953), Burmester devised a test for some of these abilities in the biological field.

Another contribution is that of Dunning (1954), who argued for the importance of *critical thinking* in science teaching and examining. In particular, Dunning considered that critical thinking should include (i) the ability to apply principles, (ii) the ability to interpret data, and (iii) the ability associated with the nature of proof. Tests were devised to measure success in each of these categories. Two of these tests have been modified for use in the present investigation.

A group research sponsored by the American Council on Education was reported by Dressel and Mayhew (1954). It was maintained that, in relation to the practical problem of test construction, the abilities of basic importance in school science should include those of recognising and stating problems, those of selecting, evaluating and applying information in relation to problems, those of recognizing and selecting hypotheses, and those of recognizing and evaluating conclusions, assumptions and generalizations.

It is apparent from this brief survey that while the elements comprising scientific ability are many, and may be stated in different ways, yet a common thread runs through most of the category lists outlined. As has been stated elsewhere (Lewis, 1965), most of the categories listed in the various studies may be ordered along an ability dimension extending from the low-level factual to the high-level evaluatory. Such an ordering too would agree with the broad divisions of the Bloom Taxonomy (Bloom, 1956). This taxonomy, in fact, provides the basis for the construction of the science tests used in the present research.

III.—MEASURES OF SCIENTIFIC ABILITY.

Objective tests of ability in each of physics, chemistry and biology were formulated from items selected—sometimes with modifications—from the Dressel and Nelson file (1958). This file, comprising a collection of objective (mostly multiple-choice) science questions categorised both as regards subject matter and in accordance with the Bloom Taxonomy, was compiled for use in American high schools. Suitable questions were sent to groups of teachers, all experienced in the teaching of O-level science, for classification according to the following scheme :

(1) Knowledge of specific facts (the recall of specific and relatively isolated pieces of information).

(2) Knowledge of general principles (theories, structures, methods, etc.) by means of which material in a particular field is organized.

(3) Comprehension—requiring an understanding of the material presented, but without necessarily relating it to other material or seeing it in all its implications.

(4) Application—abstracting from a particular situation and applying the abstraction in other ways.

(5) Analysis—identifying the constituent elements in the material presented and appreciating their interconnections.

The teachers were asked to place each question in the category to which they considered it belonged. Questions were then selected for the final tests according to these category placements (and, in particular, questions receiving an unduly wide spread of placements were rejected). A number of questions were placed in a different category, usually one nearer the factual knowledge end of the continuum, from that originally assigned to them by Dressel and Nelson. No clear-cut distinction between all the categories was possible—though, on the other hand, there was an almost unanimous agreement in distinguishing between the extremes. In particular, the overlap between categories 1 and 2 (the two knowledge categories) and also between categories 3 and 4 (comprehension and application) was considerable. It was, therefore, decided to fall back on broader groupings with tests of (i) knowledge, (ii) comprehension and application, and (iii) evaluation in each of physics, chemistry and biology. Two tests of Dunning (1954)—on the interpretation of data in the physical and biological fields—together with a test on interpretation in chemistry taken from the Dressel and Nelson file were added to the analysis questions to make up category (iii), evaluation.

The nature of the tests may be seen from the examples (taken from the physics tests) shown below. All the tests were based on subject matter in the syllabuses of the Joint Matriculation Board Ordinary Level examinations in physics, chemistry and biology. The questions are in multiple-choice form, the correct answer having to be selected from five alternatives.

EXAMPLES

Physics : *Knowledge.*

(i) Power may be measured in
A. ergs ; B. dynes ; C. foot-pounds ; D. watts ; E. foot-poundals.

(ii) A 5-lb. mass does not fall five times as fast as a 1-lb. mass of the same size and shape because
A. the earth's gravitational force is the same for both ; B. the ratio of force to mass is the same for both ; C. air friction has more effect on the 5-lb. mass ; D. when two masses have the same kinetic energy, the lighter one moves faster ; E. all masses fall at the same rate in a vacuum.

Physics : *Comprehension and Application.*

(iii) A force of 10-lb. and another of 12-lb. may be applied so that their resultant is
A. 0-lb. ; B. 1-lb. ; C. 10-lb. ; D. 24-lb. ; E. 120-lb.

(iv) If the pressure is constant and a gas has a volume of 10 cubic feet at 20°C, its volume at 40°C will be
A. 5 cubic feet ; B. between 9 and 10 cubic feet ; C between 10 and 11 cubic feet ; D. 20 cubic feet ; E. over 20 cubic feet.

Physics : *Evaluation.*

(v) A sample of coal was placed in a thin metal capsule which was then filled with oxygen and sealed. It was then immersed in an insulated vessel of water, and the contents were ignited by an electric spark. Which of the following measurements is *least* essential for determining the heat produced by the type of coal used :
A. weight of coal sample ; B. weight of water ; C. weight of capsule ; D. final temperature of the water ; E. time required for combustion.

(vi) If two objects X and Y have the same momentum, Y can have more kinetic energy than X only if it
A. weighs more than X ; B. is moving faster than X ; C. weighs the same as X ; D. is moving more slowly that X ; E. occupies a position above that of X.

IV.—DESIGN OF THE EXPERIMENT.

(i) *Factor Analysis.*

The measures of ability in physics, chemistry and biology were factor-analyzed together with objective tests of verbal, spatial and numerical ability, the general design being as outlined in Table 1. Assuming that verbal, spatial and numerical group factors are obtained, the question marks in Table 1 would correspond to the questions for which the experiment might provide answers, namely, to what extent does ability in science (and in its separate branches and categories) depend upon verbal, spatial and numerical ability. The tests of the

TABLE 1

Tests	Group Factors			
	Verbal	Numerical	Spatial	Scientific
Vocabulary	×			
Controlled Association	×			
Verbal Reasoning	×			
Language Usage	×			
Number Series		×		
Number Pattern		×		
Arithmetic		×		
General Ability (Perception) I			×	
General Ability (Perception) II			×	
Space Relations			×	
Physics: Knowledge	?	?	?	×
Physics: Comprehension and Application	?	?	?	×
Physics: Evaluation	?	?	?	×
Chemistry: Knowledge	?	?	?	×
Chemistry: Comprehension and Application	?	?	?	×
Chemistry: Evaluation	?	?	?	×
Biology: Knowledge	?	?	?	×
Biology: Comprehension and Application	?	?	?	×
Biology: Evaluation	?	?	?	×

battery, however, can also be ordered along a continuum relating to their dependence on specific instruction and formally acquired skills. An alternative grouping of tests, one based on an achievement-aptitude dichotomy, is shown in Table 2. The question marks appearing alongside the tests of comprehension and application, and evaluation in physics, chemistry and biology, would then

TABLE 2

Tests	Achievement	Aptitude
Vocabulary	×	
Controlled Association		×
Verbal Reasoning		×
Language Usage	×	
Number Series		×
Number Pattern		×
Arithmetic	×	
General Ability (Perceptual) I		×
General Ability (Perceptual) II		×
Space Relations		×
Physics: Knowledge	×	
Physics: Comprehension and Application	?	?
Physics: Evaluation	?	?
Chemistry: Knowledge	×	
Chemistry: Comprehension and Application	?	?
Chemistry: Evaluation	?	?
Biology: Knowledge	×	
Biology: Comprehension and Application	?	?
Biology: Evaluation	?	?

correspond to the questions of to what extent these higher-level aspects of ability in school science depend upon (i) general achievement and (ii) a scholastic aptitude relatively free from formally acquired skills. Whether a factor of scholastic aptitude, covering tests of widely differing contents, will emerge cannot be known beforehand. It is prudent, however, not to rely on a method of analysis, which depends on a prior grouping of tests. The method of varimax simple structure was in fact employed.

(ii) *Tests.*

The science tests have been described in Section III. The verbal tests used were:

(i) Vocabulary—an objective test of vocabulary constructed by the writer.
(ii) Controlled Association (Thurstone, 1938).
(iii) Verbal Reasoning—sub-test III from the General Ability Test, Verbal, of the Differential Test Battery (Morrisby, 1955).
(iv) Language Usage—a test designed to measure the ability to distinguish between the correct and incorrect use of language (similar in form to Part II of the Language Usage Test of the Differential Aptitude Tests (Bennett, Seashore and Weisman, 1959).

The numerical tests were:

(v) Number Series } Sub-tests II and III of the Differential Test Battery
(vi) Number Pattern } (Morrisby, 1955).
(vii) Arithmetic—a multiple-choice test of numerical problems constructed by the writer.

The spatial tests were:

(viii) General Ability (Perceptual) I } Sub-tests II and III from the General
(ix) General Ability (Perceptual) II } Ability Test, Perceptual, of the Differential Test Battery (Morrisby, 1955).
(x) Space Relations, Form A (Bennett, Seashore and Weisman, 1959).

(iii) *Groups.*

Testing was carried out in five schools in the Manchester area, all the testees being boys who had completed O-level courses in physics, chemistry and biology. The numbers involved, together with the ages of each group, are shown in Table 3.

TABLE 3

THE GROUPS TESTED.

School	Numbers	Mean Age (yrs. : mths)	Age Range (yrs. : mths)
1	41	15 : 10	14 · 9—16 : 11
2	34	16 : 7	15 : 7—17 : 8
3	39	15 : 6	14 : 11—16 : 0
4	39	15 : 8	14 . 8—16 : 6
5	38	15 : 9	14 : 8—16 : 9

V.—ANALYSIS OF RESULTS.

Appreciable and statistically significant differences between the school groups were found for the scores on the science tests and also for the scores on most of the other tests. Correlations between tests were, therefore, derived separately for each group (see Lindquist, 1940, 219-228) and averaged via Fisher's z transformation. The averaged correlations were factor-analyzed, four significant factors being obtained. These were then rotated to simple structure in accordance with the varimax criterion (Kaiser, 1958). The factor loadings are shown in Table 4.

TABLE 4

VARIMAX FACTOR LOADINGS.

(Decimal points omitted in body of table.)

Test	Factor				Communality
	I	II	III	IV	
Vocabulary	301	148	—238	588	0·515
Controlled Association	—003	094	222	719	0·575
Verbal Reasoning	—084	289	101	746	0·657
Language Usage	284	158	—062	492	0·352
Number Series	014	385	724	048	0·675
Number Pattern	158	201	798	—031	0·703
Arithmetic	356	065	729	120	0·677
Space Relations	143	611	205	271	0·509
General Ability: Perceptual I	043	755	096	142	0·601
General Ability: Perceptual II	171	763	197	—018	0·651
Physics: Knowledge	677	230	100	121	0·536
Physics: Comprehension and Application	689	248	213	—048	0·584
Physics: Evaluation	360	410	326	089	0·412
Chemistry: Knowledge	717	—164	077	196	0·585
Chemistry: Comprehension and Application	695	001	215	—042	0·531
Chemistry: Evaluation	361	301	145	274	0·317
Biology: Knowledge	773	097	—069	—003	0·612
Biology: Comprehension and Application	747	114	—031	077	0·578
Biology: Evaluation	404	471	160	—040	0·421
Sum of Squares	3·891	2·502	2·169	1·921	10·483
Contribution to Variance (per cent.)	20·5	13·2	11·4	10·1	55·2

We see that Factor I is a factor with substantial loadings on the science tests, as well as appreciable loadings on Vocabulary, Language Usage and Arithmetic. A point of some importance is that the loadings on the science tests are not uniformly high, the loadings on the three tests in the evaluation category being noticeably lower than those on the six tests in the categories of knowledge and of comprehension and application. Clearly, the factor is not one covering all aspects of scientific ability in equal measure. Rather it is one in which the higher-level evaluatory aspect is of markedly less importance. Since, too, the factor has appreciable loadings on such essentially 'scholastic' tests as Vocabulary, Language Usage and Arithmetic (and with markedly smaller loadings on the tests of reasoning) it should be regarded as a factor of scholastic, including scientific, achievement. The factor is the most influential of the four, contributing over 20 per cent. to the total variance.

Factor II is a factor defined to a considerable extent by the high loadings on the three spatial tests. It also has moderate loadings on the three science tests of evaluation—the loadings on the remaining science tests being appreciably smaller. As far as the number and verbal tests are concerned, the highest loadings are on Verbal Reasoning and Number Series—two tests in the aptitude classification of Table 2. The factor may best be regarded, therefore, as a factor of scholastic aptitude or reasoning.

Factor III is clearly a factor of numerical ability, the loadings on the three number tests predominating over all else. The only other loading of any consequence in fact is that on the test of evaluation in physics (·326).

Factor IV has high or moderately high loadings on all the verbal tests, and may be interpreted as a straight-forward verbal factor. Its loadings on all the science tests are clearly negligible, except for that on the test of evaluation in chemistry (·274).

In summary then the results of the analysis yield a factor of scholastic achievement accounting in substantial measure for the knowledge and the comprehension and application categories of all three school sciences, an aptitude factor accounting in part for the evaluation category of all the three sciences, together with verbal and numerical factors which have little overlap with science—except possibly for some dependence of the evaluation category of physics on the number factor. It may be added, too, that although only results from correlations averaged over all the schools are provided in this paper, a fundamentally similar factorial pattern was obtained from most of the separate school groups. Details for four of the five groups are provided in an unpublished thesis (Lewis, 1966)

VI.—DISCUSSION.

The science tests used in this experiment are comprehensive in that they attempt to take account not merely of what is actually taught and examined as science in Ordinary Level courses of the General Certificate of Education, but also what informed opinion considers *should* be taught. The review of the relevant literature, together with the category placements of questions by the teachers co-operating in the enquiry, suggests that the structuring of the tests into three broad categories—those of basic knowledge, comprehension and application, and evaluation—represent a consensus of what scientific ability at this level necessarily involves

It follows too that the science tests of this experiment differ among themselves in two ways. There are the differences between subjects (physics, chemistry and biology) across categories, and the differences between categories across subjects. It should be noted, too, that the same form of the tests (multi-choice) used throughout enables a comparison of the two differences to be made without distortion from a formal factor. The results show that it is the differences between categories which are by far the more important. Indeed only minor differences between subjects were forthcoming. The conclusions derived from the experiment apply to all three of physics, chemistry and biology alike.

The important difference is that found between the categories of knowledge and of comprehension and application on the one hand, and the category of evaluation on the other. While the instruction and training received by the testees in their O-level science courses is such that their performance on the tests of comprehension and application is basically (i.e., factorially) similar to that on the tests of factual knowledge, their performance on the tests of evaluation is noticeably different. As far as the evaluation of scientific data is concerned, the formal training received appears to be less relevant, the performance on the tests in this category being influenced to about an equal extent by general reasoning. The school courses have, of course, been concerned first and foremost with preparation for the O-level examinations, examinations in which it has been noted the emphasis is placed on the lower-level aspects of scientific ability. It may well be therefore that the pattern of factor loadings on the science tests shown in Table 4 reflects an insufficient training in the higher-level aspects of scientific ability as measured in particular by the evaluation tests of this research. It would be interesting to see if the same factorial pattern occurs in the future if, as appears likely, the present mode of examining science in schools undergoes a change.

A further point is that although distinctive numerical and verbal group factors were isolated by the factorial analysis, these factors are not involved in any marked degree in the performance on the science tests. Group factors of content alone do not assume a major role.

Scientific ability has been measured in this experiment solely by objective, multiple-choice tests, and it might be fairly objected the measurement of certain aspects of ability desirable, if not essential, in the make-up of scientists has thereby been excluded. Certainly, the category of synthesis in the Bloom Taxonomy has not been directly tested, and the omission too of any assessment of practical ability springs to mind. Few would argue, however, that objective tests of the type used should completely replace the existing forms of examination. The advantages of supplementing present examination techniques by objective tests, and particularly by objective tests of the non-factual, higher-level categories of scientific ability, are nevertheless, strongly suggested from the results of this research.

VII.—References.

Bennett, G. K., Seashore, H. G., and Weisman, A. G. (1959). *Manual for the Differential Aptitude Tests.* New York : The Psychological Corporation.

Berridge, A. R. (1947). *A Factorial Analysis of Ability in School Physics.* Unpublished M.A. Thesis, University of London.

Bloom, B. S. (1953). *Taxonomy of Educational Objectives : The Classification of Educational Goals.* Handbook 1 : Cognitive Doman. London : Longmans.

Burmester, M. A. (1952). Behaviour involved in the critical aspects of scientific thinking. *Science Education*, 36, 5, 259-263.

Burmester, M. A. (1953). The construction and validation of a test to measure some of the inductive aspects of scientific thinking. *Science Education.* 37, 2, 131-140.

Crossland, R. W., and Amos, R. (1961). What do ' O ' level examinations in biology test? *Biology and Human Affairs*, 26, 2, 1-4.

Crossland, R. W., and Amos, R. (1965). Towards new ' O ' level examinations in biology. *Biology and Human Affairs*, 30, 3, 1-5.

Dressel, P. K., and Mayhew, L. B. (1954). Objectives in science, Chapter V in *General Education : Explorations in Evaluation.* Washington D.C. : American Council on Education.

Dressel, P. L., and Nelson, C. H. (1958). *Questions and Problems in Science.* Test Item Folio, No. 1. Princeton : Educational Testing Services.

Dunning, G. M. (1954). Evaluation of critical thinking. *Science Education*, 38, 3, 191-211.

Ellis, A. G. (1948). *The Relations between Abilities in Grammar School Subjects and their bearing upon School Organization.* Unpublished M.Ed. Thesis, University of Durham.

Engelhardt, M. D., and Lewis, H. B. (1941). An attempt to measure scientific thinking. *Educational and Psychological Measurement*, 1, 289-294.

FRUTCHLEY, F. P., and TYLER, R. W. (1937). Examinations in the natural sciences, Chapter V in *The Construction and Use of Achievement Examinations* (edited by Hawkes, H. E., Lindquist, E. F., and Mann, C. R.). London: Harrap.

GULBENKIAN FOUNDATION REPORT (1959). *Report of an Enquiry into the suitability of the General Certificate of Education Advanced Level Syllabuses in Science as a Preparation for Direct Entry into First Degree Courses in the Faculty of Science.* University of Birmingham.

HEIL, L. M., KAMBLY, P. E., MAINARDI, M., and WEISMAN, L. (1946). The measurement of understanding in science, Chapter VI in *The Forty-fifth Yearbook of the National Society for the Study of Education*, Part I (editor, Henry, N. B.). Chicago: University of Chicago Press.

JAMES, E. (1951). *An Investigation into the Factors entering into the Ability to learn Biology at School Certificate stage.* Unpublished M.A. Thesis, University of London.

JOG, R. N. (1955). *A Study of the Factors underlying the ability of Grammar School Boys, at the School Certificate stage, to learn Physics.* Unpublished M.A. Thesis, University of London.

KAISER, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23, 187-200.

KEESLAR, O. (1945). The elements of scientific method. *Science Education*, 29, 5, 273-278.

LAMB, F. C. (1954). *A Factorial Analysis of Ability in School Biology.* Unpublished M.Ed. Thesis, University of Nottingham.

LEWIS, D. G. (1961). Group factors of attainment in grammar school subjects. *Brit. J. Educ. Psychol.*, XXXI, 241-248.

LEWIS, D. G. (1965). Objectives in the teaching of science. *Educ. Res.*, VII, 3, 186-199.

LEWIS, D. G. (1966). *An Experimental Study of Ability in Elementary Science.* Unpublished Ph.D. Thesis, University of Manchester.

LINDQUIST, E. F. (1940). *Statistical Analysis in Educational Research.* Houghton Mifflin Co.

MORRISBY, J. R. (1955). *Differential Test Battery.* National Foundation for Educational Research in England and Wales.

ORMISTON, M. (1939). The bearing of general and special abilities upon scholastic success at the beginning and end of a secondary school career. *Brit. J. Educ. Psychol.*, IX, 164-173 and 213-223.

PAWLEY, G. C. (1937). *A Factor Anaylsis of the Abilities involved in the learning of School Chemistry.* Unpublished M.A. Thesis, University of London.

POLLARD, H. A. (1942). *The Solution of Concrete Problems.* Unpublished M.A. Thesis, University of Birmingham.

SCIENCE MASTERS' ASSOCIATION (1938). *The Teaching of General Science*, Part II. London: Murray.

TEICHMAN, L. (1944). The ability of science students to make conclusions. *Science Education*, 28, 5, 268-279.

THURSTONE, L. L. (1938). *Primary Mental Abilities.* Chicago: University of Chicago Press.

WRIGLEY, J. (1956). *The Factorial Nature of Ability in Elementary Mathematics.* Unpublished Ph.D. Thesis, Queen's University of Belfast.

(*Manuscript received 10th November, 1966*)

VALUE PATTERNS OF ENGINEERING, LAW, MEDICAL AND TEACHER-TRAINING STUDENTS IN INDIA

By S. K. PAL

(*University of Allahabad, India*)

SUMMARY. The study examined the value patterns of engineering, law, medical and teacher-training students in Indian colleges. Fifty students in each group were administered the Allport-Vernon-Lindzey Study of Values. When the t-test was applied to the strength of six value variables within groups, a clear hierarchy of values emerged in each group. In the case of engineering students, economic value was the highest value variable while religious value was the lowest. In the case of law students, political value emerged as the highest value variable followed by economic value. In the case of medical students, theoretical value emerged as the highest value variable followed by social value, while aesthetic value emerged as the lowest value variable. In the case of teacher-training students, political value was the first followed by theoretical value.

I.—INTRODUCTION.

A review of the literature tends to support the point of view that psychological values of students in important professions have not received a thorough exploration. Some studies, however, have demonstrated that students in selected professions show distinguishing patterns of values. Gee (1960) described the value patterns of medical students in terms of highest and lowest scores. The medical students were characterized by high theoretical value and low economic value. The political value was scored second high and social value was one of the two lowest. Allport, Vernon and Lindzey (1960) reported that the engineering students placed high esteem on the theoretical, economic and political values in order named. The religious value was ranked fourth and the lowest assessment was for the aesthetic scale. The present study was designed to discover the value patterns of engineering, law, medical and teacher-training students in India.

II.—METHOD.

The present study employed the Allport-Vernon-Lindzey Study of Values which measures six values on a relative rather than absolute basis. The six values are: theoretical, economic, aesthetic, social, political and religious. The Study of Values was administered to groups of students consisting of fifty fourth-year engineering students, fifty final-year law students, fifty fourth-year medical students and fifty final-year teacher-training students. All were males. Special care was exercised in selecting the subjects to insure that each had genuine interest in, and expressed clear indication of commitment to, the profession. The value variables of each group were ranked on the basis of mean scores. The *t* test was applied to the strength of six value variables within each group.

III.—RESULTS.

The pattern of highest and lowest scores of the four groups was found by ranking the six value variables from highest to lowest for each group on the basis of average score. The relative positions of different values are given in Table 1.

TABLE 1

RANK ORDER OF VALUES FOR ENGINEERING, LAW, MEDICAL AND TEACHER-TRAINING STUDENTS.

Ranks	Engineering	Law	Medical	Teacher
I	Economic (45·76)	Political (44·64)	Theoretical (44·30)	Political (42·62)
II	Theoretical (40·26)	Economic (41·18)	Social (43·12)	Theoretical (42·38)
III	Political (40·12)	Theoretical (40·08)	Economic (40·70)	Economic (39·82)
IV	Social (39·26)	Social (38·98)	Political (40·06)	Aesthetic (38·90)
V	Aesthetic (38·60)	Religious (37·58)	Religious (37·40)	Social (38·74)
VI	Religious (36·14)	Aesthetic (37·42)	Aesthetic (34·34)	Religious (37·54)

When the *t* test was applied to the strength of six value variables within groups a clear hierarchy of values emerged in each group.

TABLE 2

'*t*' MATRIX SHOWING INTRA VALUE DIFFERENCES OF ENGINEERING STUDENTS.

	Theo	Eco	Aes	Soc	Pol	Rel
Theo						
Eco	4·04†					
Aes	1·30	5·68†				
Soc	0·91	4·85†	0·48			
Pol	0·14	5·17†	1·24	0·72		
Rel	3·35†	7·19†	1·66	2·58*	3·11†	

* Significant beyond 0·05 level.
† Significant beyond 0·01 level.

Table 2 shows that, for engineering students, the economic value differs significantly from all other values and religious from all the other except aesthetic Other values do not differ significantly from one another. Economic value thus emerges as the highest single value variable, while religious as the lowest. Thus, economic and religious values occupy, respectively, the highest and lowest places in the group of engineering students, while the other four form a separate cluster with no significant difference among them.

TABLE 3

'*t*' MATRIX SHOWING INTRA VALUE DIFFERENCES OF LAW STUDENTS.

	Theo	Eco	Aes	Soc	Pol	Rel
Theo						
Eco	0·96					
Aes	1·69	2·63*				
Soc	0·80	1·50	0·94			
Pol	3·83†	2·62*	4·94†	3·47†		
Rel	1·67	2·31*	0·10	0·95	4·41†	

It will be observed that, in law students, political value differs significantly from all other values, while economic value differs significantly from aesthetic, political and religious values. There is no significant difference between the relative strengths of theoretical, aesthetic, social and religious values. Thus, two clusters of values emerge : (1) political and economic, and (2) theoretical, aesthetic, social and religious. Political value emerges as the highest value variable, followed by economic value.

TABLE 4

'*t*' MATRIX SHOWING INTRA VALUE DIFFERENCES OF MEDICAL STUDENTS.

	Theo	Eco	Aes	Soc	Pol	Rel
Theo						
Eco	4·00†					
Aes	7·78†	5·39†				
Soc	1·01	1·60	5·42†			
Pol	4·51†	0·52	4·97†	2·73†		
Rel	5·48†	2·58*	2·20*	4·00†	2·23*	

The '*t*' matrix of medical students (Table 4) shows that theoretical and social values do not differ significantly and they are both on the high side. Similarly political and economic values also do not differ significantly from one another, though these students are definitely low on these values, as compared to their theoretical and social values. They are lowest on religious and aesthetic values which between themselves differ significantly at 0·05 level. The aesthetic value is the lowest among medical students. In this way three clusters of values emerge : (1) theoretical and social, (2) political and economic, and (3) religious and aesthetic. The theoretical value emerges as the highest followed by social value ; aesthetic value emerges as the lowest.

TABLE 5

'*t*' MATRIX SHOWING INTRA VALUE DIFFERENCES OF TEACHER-TRAINING STUDENTS.

	Theo	Eco	Aes	Soc	Pol	Rel
Theo						
Eco	1·95					
Aes	2·17*	0·56				
Soc	2·62*	0·79	0·14			
Pol	0·19	2·29*	2·95†	3·13†		
Rel	3·10†	1·50	1·03	1·06	3·76†	

Table 5 shows that in the case of teacher-training students, the political value differs significantly from four others, namely, religious, social, aesthetic and economic and theoretical from religious, social and aesthetic. Political value emerges as the highest value variable followed by theoretical. The other four values, religious, social, aesthetic and economic, form a separate cluster with no significant difference among them.

IV.—DISCUSSION.

The results indicate that the students preparing for four different professions under consideration show some distinguishing patterns of values. In the case of engineering students, economic value emerged as the highest single value variable, while religious was the lowest. This interest in the practical and useful seems to be compatible with the nature of work which they have to perform.

In the case of law students, political value had the highest rank, followed by economic value. This high esteem on the political value would seem to indicate that the law students gain satisfaction from power, prestige, and in having some control over others. Great importance given to economic value may also show that they are interested in that which is useful and practical.

In the case of medical students, theoretical value was first followed by social, while aesthetic value was the lowest variable. High importance attached to the theoretical value indicates that they probably gain satisfaction from ordering and systematizing knowledge. Since medical students are constantly reminded to keep the ideal of service to humanity as their prime consideration, it is not perhaps surprising that they score highly on the social value. The low value assessment for the aesthetic approach, perhaps, indicates a lack of interest in form, harmony and such artistic aspects of life.

In the case of teacher-training students, political value had first position, then theoretical. This suggests that these students may be interested in power, prestige and in having some control over others. If further research should result in similar findings, we may need to revise some notions about the personality motivation of future teachers. Teacher-training students also placed great importance on the theoretical value showing that they are oriented to a ' cognitive approach ' consistent with the nature of the work of a teacher.

Findings in general lead to the conclusion that the students preparing for the professions of engineering, law, medicine and teaching do possess distinguishing patterns of values. Further investigation is needed to establish whether these programmes attract students with such value patterns or such patterns develop in the course of training.

V.—REFERENCES.

ALLPORT, G. W., VERNON, P E., and LINDZEY, G. (1960). *Manual of Directions : Study of Values.* Boston · Houghton Mifflin Co., third edit., p. 14.

GEE, H. H. (1960). Differential characteristics of student bodies-implications for the study of medical education, in *Selection and Educational Differentiation, Report of a Conference*, May 25-27, 1959. Berkeley : Field Service Center and Center for the Study of Higher Education, pp. 140-46.

(*Manuscript received* 17*th November*, 1966)

RESEARCH NOTES

SOME EVIDENCE ON THE VALIDITY OF THE JUNIOR MAUDSLEY PERSONALITY INVENTORY

By A. H. MEHRYAR
(*Centre for Teacher Training and Educational Research, Tehran, Iran.**)

I.—PLAN OF THE INVESTIGATION.

The present paper reports some data regarding the predictive and concurrent validities of the Junior Maudsley PersonalityInventory as reflected in its correlation with the results of a sociometric test and a peer-nomination technique.

Seventy-nine children, forty-two boys and thirty-seven girls, between the ages of 12 and 14 years 4 months, median age 13 years 8 months, comprising the population of three classes in an East London secondary modern school, were given the following tests:

(i) Junior Maudsley Personality Inventory.

(ii) A sociometric questionnaire in which the children were asked to choose:
Someone to work with in a group in the English class.
Someone to work with in a group in the Maths. class.
Someone to spend your free time with.
The most liked person.
The most (and least) popular member of the class.
Someone to act as your representative in an important conference of student leaders.
A close friend.

The number of choices on each criterion was limited to five from within the class and the subjects were asked separately for their choices and rejections on the first five items.

(iii) A twenty-item Guess Who Test (G.W.T.), with up to five choices on each item.

TABLE 1

THE TEXT OF THE GUESS WHO TEST.

No.	
1	Someone whom you regard as very good at sports.
2	Someone who takes life easily, who does not worry and get upset.
3	Someone very good at English.
4	Someone who is very fond of people, who mixes with other people easily and has a lot of friends.
5	Someone whose feelings are easily hurt and who is quickly offended.
6	Someone you can rely on, who is very sincere and keeps his or her word.
7	Someone who does not care much about other people and is concerned only with his or her own interests.
8	Someone very co-operative and friendly who is always willing to help others.
9	Someone who is very ill-tempered and quarrelsome, who loses his or her temper very quickly and gets irritated easily.
10	Someone with a very good sense of humour, who is very fond of cracking jokes and telling stories to his or her friends.
11	Someone who worries a lot about his or her failures, mistakes and blunders.
12	Someone who is shy, who does not like being watched or criticised by people.
13	Someone who is bossy and domineering, who likes to order other people about.
14	Someone who is not interested in being with other people.
15	Someone with a strong influence on others; someone whose opinions are generally accepted by most of his friends.
16	Someone with a very low opinion of herself or himself; this person never thinks he or she does anything well.
17	Someone very good at Maths.
18	Someone who always takes the lead in group activities.
19	Someone who frequently gets stuck, grows disheartened and so never finishes what he has started.
20	Someone who feels sure he can cope, who just gets on with a job and gets it done.

---

* Now at Pahlavi University, Shiraz, Iran.

Three of these items (Nos. 1, 3, 17) dealt with high achievement in Sports, English, and Maths., respectively. The remaining seventeen items were selected in a way to cover the main traits of personality underlying the dimensions of extraversion and neuroticism as summarised in Eysenck's (1964) circle diagram.

The positive pole of the extraversion dimension is assumed to be covered by the following items: 2 (Easy-going), 4 (Sociable), 10 (Good sense of humour), 15 (Social influence) and 18 (Leadership). The negative pole of extraversion is represented by items 12 (Shy) and 14 (Unsociable). Items with an apparent bearing on Neuroticism are: 5 (Easily hurt), 7 (Selfish), 9 (Ill-tempered), 11 (Worried), 13 (Bossy and domineering), 16 (Lacking in self-confidence), and 19 (Lacking in perseverance). Those with an apparently opposite implication are 6 (Reliable), 8 (Co-operative) and 20 (Self-confidence and perseverance). For reasons outside the scope of this report, the G.W.T. had to be administered roughly four mohths after the administration of the first two tests.

In view of the variation in the number of choices and nominations exchanged within each class, it was decided to convert the sociometric and Guess Who Test scores into percentages of those made on each item. Two further items representing the sum of raw scores on criteria 1–4 of the sociometric questionnaire and items 1–20 of the G.W.T. were added and product moment correlation coefficients were computed between each of these scores and the subjects' extraversion and neuroticism scores on the J.M.P.I.

II.—RESULTS.

The means and standard deviations of the group on the J.M.P.I. were: Extraversion: 12·10 ± 3·40; Neuroticism: 7·40 ± 3·09, r between the two being —·13. The product moment correlation coefficients between each of these two variables and the various criteria of the sociometric questionnaire are listed in Table 2. This reveals a consistently significant and positive—although not very high—correlation between extraversion as measured by the J.M.P.I. and choice status on all the criteria used.

TABLE 2

CORRELATIONS BETWEEN THE J.M.P.I. VARIABLES AND THE SOCIOMETRIC CRITERIA OF CHOICE AND REJECTION.

	Sociometric Criteria													
	Choice								Rejection					
J.M.P.I.	1	2	3	4	5	6	7	1–7	1	2	3	4	5	1–4
Extraversion	·30	·24	·40	·39	·44	·25	·33	·37	·05	–·01	–·11	–·07	–·38	–·05
Neuroticism	·12	·06	·11	·03	·01	–·06	·09	·06	–·09	–·12	–·12	–·12	·05	–·15

NOTE.—With a d.f. = 77, r = ·22, p < ·05, r = ·28, p < ·01 (one-tailed test)

Thus, those scoring high on the extraversion scale would seem to have been rated as more popular and likeable and to have been chosen as 'free-time companions,' 'friends,' 'co-workers,' and 'leaders' more often than those scoring low on the scale. The negative correlations between extraversion and being rejected on the same criteria, however, do not reach statistical significance except in the case of 'popularity.'

The correlations between the J.M.P.I. and the G.W.T. scores are summarised in Table 3. Here, significant positive correlations are found between extraversion and items 2 (easy-going), 4 (sociable), 10 (sense of humour), 13 (bossy and domineering), 15 (social influence) and 19 (lack of confidence ?). Near significant correlations (i.e., $\cdot17$ to $\cdot20$, $\cdot05 < p < \cdot1$) are also obtained with items 1 (good at sports), 3 (good at English), 17 (good at Maths.),18 (Leadership), and 20 (Confidence and perseverance). All these correlations—with the possible exception of that of item 19—are in the generally expected direction. So also are the negative—although non-significant—correlations with items 11 (Worrier) and 14 (Unsociable). The correlation between the sum of the G.W.T. scores and extraversion is also moderately high and significant.

TABLE 3

Correlations Between the J.M.P.I. Variables and the Guess Who Test.

	Guess Who Test Items																				
J.M.P.I. . .	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	1–20
E	·19	·29	·20	·30	·02	·07	·03	·09	·11	·45	–·14	–·05	·29	–·14	·42	·09	·17	·20	·22	·18	·37
N	–·06	·04	–·16	·06	·22	·02	·02	·02	·14	·06	–·07	–·01	·05	–·18	–·03	·06	–·22	·02	·00	–·26	–·07

Note.—With d.f. $=77$, $r=\cdot22$, $p<\cdot05$; $r=\cdot28$, $p<\cdot01$ (One-tailed test of significance) .

The Neuroticism scale shows a significant correlation only with items 20 (Self-confidence), 17 (Achievement in Maths.) and 5 (Easily hurt). Near significant correlations are also found with items 14 (Unsociable), 3 (Achievement in English) and perhaps 9 (Ill-tempered and quarrelsome). Thus, those who have admitted a greater number of neurotic tendencies on the J.M.P.I. have tended to be rated as lacking in self-confidence, poor at Maths. and English, and highly irritable and quarrelsome, by their classmates. These are all more or less in accord with common-sense expectation. The negative—although barely significant—correlation between neuroticism and Unsociability (item 14) is, however, difficult to explain.

III.—Discussion.

Several studies have indicated the predominance of a ' sociability ' component in the factorial structure of extraversion (Eysenck and Eysenck, 1963 ; Hallworth, 1964). It is to be expected that the self-assessed extraversion on the J.M.P.I. should be reflected in such objective measures of sociability as covered by the seven criteria of sociometric choice included in this study. The relatively low size of these correlations may be due to the crudeness of the sociometric technique as a criterion as well as the low reliability of the test itself. Yet, it is encouraging to note that other investigations into the relationship between extraversion and sociometric status have found similarly low correlations (Mann, 1959).

The absence of any correlation between extraversion and rejection status, except on the criterion of popularity, is due to the structural properties and the more or less haphazard nature of sociometric rejection process. Far from being the logical opposite of choice, rejection forms a category of its own and, unlike the choice process, does not seem to show any significant deviation from the pattern expected on the basis of a purely random allocation of votes (Mehryar, 1965).

Tha extraversion as measured by the J.M.P.I. is also significantly correlated with peer-ratings on such characteristics as sense of humour, social influence, sociability, easy-going, bossiness, etc., further confirms the validity of the self-assesssed variable. Again, the relatively low sizes of the correlations may be due to the crudity of the G.W.T. and the fairly long interval between the administration of the two tests. If some of the items correlating with extraversion do not seem to have an obvious bearing on the dimension as generally understood (e.g., item 19), the reason must be sought in the subjective implications of these characterisations for the children concerned as well as the influence of a general ' halo effect ' running through rating scales of this kind. The latter factor, along with a tendency to make fewer nominations on the socially undesirable items of the G.W.T. (Mehryar, 1967) is also partially responsible for the significant correlation between extraversion and the sum of the G.W.T. scores.

The failure of the neuroticism scale to correlate significantly with any of the sociometric measures may indicate the irrelevance of this dimension for interpersonal interactions in a normal sample. This is in agreement with the findings of other studies. Mann (1959), reviewing eighteen independent studies, found the median correlation between ' adjustment '—the positive pole of neuroticism—and sociometric popularity to be close to ·10. No single measure of adjustment was convincingly related to popularity except an opinion survey designed to measure satisfactory personal habits. The fact that neuroticism as measured by the J.M.P.I. correlated significantly with peer-ratings on such obviously relevant traits as being easily hurt and lacking in self-confidence provides some evidence for the validity of the scale.

ACKNOWLEDGEMENT.—Thanks are due to Mrs. D. M. Penfold and Mr. H. E. O. James who supervised this work; to the staff and pupils of the Goodall Secondary School, Leighton, where this study was carried out; and to the Iranian Ministry of Education for their financial support. I am indebted to Dr. Keith Hope for permission to use his computer programme for the analysis of this part of the data.

V.—REFERENCES.

CALLARD, P. M., and GOODFELLOW, C. L., (1962). Neuroticism and extraversion in school boys, as measured by the Junior Maudsley Personality Inventory. *Brit. J. Educ. Psychol.*, 32, 3, 241.

COSTELLO, C. G., and BRACHMAN, H. M. (1962). Cultural and sex differences in extraversion and Neuroticism reflected in responses to a children's personality inventory. *Brit. J. Educ. Psychol.*, 32, 3, 254.

EYSENCK, H. J. (1964). Principles and methods of personality description, classification and diagnosis. *Brit. J. Psychol.*, 55, 3, 284.

EYSENCK, S. B. G. and EYSENCK, H. J. (1963). On the dual nature of extraversion. *Brit. J. Soc. Clin. Psychol.*, 2, 1, 46

FURNEAUX, W. D., and GIBSON, H. B. (1961). A children's personality inventory designed to measure neuroticism and extraversion. *Brit. J. Educ. Psychol.*, 31, 204-207.

HALLWORTH, H. J. (1964). Personality ratings of adolescents . a study in a comprehensive school. *Brit. J. Educ. Psychol.*, 34, 2, 171.

MANN, R. D. (1959). A review of the relationships between personality and performance in small groups. *Psychol. Bull.*, 56, 5, 241.

MEHRYAR, A. H. (1965). *Social Perception : An Experimental Study.* Unpublished Ph.D. thesis. University of London Library.

MEHRYAR, A. H. (1967). *Social Perception and Social Desirability. Psychological Reports.* In press.

(*Manuscript received 2nd February,* 1967)

THE AURAL ANALYSIS OF SOUNDS

By CLIFFORD CARVER
(*Manchester College of Education*)

I.—Introduction.

During a wider research (Carver, 1966) it was hypothesised that defective aural perception—specifically defective aural analysis of *sounds* into component parts—was the predominant factor in a child's inability to recognise words in visual form.

Without the ability to discriminate between 'sounds,' spoken language would be impossible. This ability to discriminate presupposes a neurological ability to analyse. It is reasonable to assume that this aural analysis is gradually acquired and consequently that children differ in this ability. It is a reasonable hypothesis that some sounds present more difficulty in analysis than others.

II.—Problem.

To confirm experimentally whether 'sounds' differed in their vulnerability to analysis.

III.—Method.

A sample of fifty-four boys and one-hundred girls, being a mixture of normal 7-year-olds and older retarded readers, were given seven words *orally*, and asked to identify the initial sound of each. Each word began with a consonant, which was followed by a vowel. (COME POKE FAT BACK GIVE SUPPER DAY.) The answering was done by ringing a letter selected from a multiple choice series of letters, provided on an answering sheet. For example : item 1 COME (given orally) : g *c* t w s y.

The same children were subsequently given the same test, using an identical answering sheet, but with seven new words substituted in place of the original ones. The answers would be the same since the new words began with the same initial consonants. However, instead of each consonant being followed by a vowel, *it was now followed by a second consonant.* These words were : CLIMB, PLEASE, FRIEND, BLADDER, GRIPPING, STRAIGHT, DREAM.

The expectation, based on previous research, was that the children would find it more difficult to hear, that is analyse, the initial consonant which was followed by a second consonant, than when followed by a vowel.

The first test enabled the child to establish his own *visual letter knowledge.* Any difference between the two tests must be attributed to the difference in the words used. The difference is assumed here to be the insertion of the second confounding consonant. It might be noted that the practice given in carrying out the first test *ought* to cause the second test to be 'easier' if no actual difference existed between the two sets of words.

The total errors made on the first set (consonant/vowel) were calculated (110 errors), as were those made on the second set (187 errors on the consonant/consonant words). This was a total increase of 77 errors.

The boys *increase* in errors was 59 ; the girls increase was 18. These increases were translated into proportions (boys, ·639 ; girls ·361).

IV.—Results.

The total errors made on the second test were found to be very significantly higher than those made on the first test ($X^2 = 23{\cdot}14$, $p < {\cdot}001$).

The proportionate *increase* made by the boys was significantly greater than that made by the girls ($\bar{z} = 3{\cdot}271$, $p \simeq {\cdot}001$).

Since the results were in the expected direction, one could reasonably have used a 'one tail' test thus raising the significance of the differences.

V.—Conclusions.

That, in this sample, words—as sounds—were shewn to differ in their vulnerability to analysis.

That the boys found significantly greater difficulty than did the girls, when the aural difficulty of the words was increased.

Two reservations should be made. Firstly, that the difference between the two sets of words could be due to an unsuspected difference other than the insertion of the second consonant. Secondly, that it could be the retarded readers who were contributing predominantly to the results.

Implications.—A spoken word is a series of fluctuating 'sounds.' We attach an abstract visual symbol to a part of such a word, on the assumption that the 'sound' it represents, can be heard as a stable or consistent aural entity. If a child is unable to detect this 'sound,' his ability to learn the appropriate *visual symbol* would be profoundly affected, since he would lack the rich associations, in aural form, to which he can attach the visual letter. It would, therefore, be a mistake to assume that inability to learn 'letters' is a visual defect.

One could conclude that the teaching of word recognition should have a strong aural rather than visual basis.

Carver, C. (1966). *Left Dominance and Reading Difficulties.* M.Ed. Thesis (Manchester University), Chap. XVI.

(*Manuscript received 7th February,* 1967)

AN OBLIQUE FACTORIZATION OF WALLACH AND KOGAN'S 'CREATIVITY' CORRELATIONS

By JAMES WARD
(*Manchester University*)

Summary. Tables of intercorrelations of creativity and 'g' tests given by Wallach and Kogan, were factorized by means of the Promax oblique method. Four factors were produced. Factor I was identified with school attainment, having high loading on all but one of the 'g' tests. Factor II was clearly a creativity factor, but showed a low correlation with Factor I. Factor III was associated with the number of responses given. Factor IV emerged in the oblique analysis as a general factor negatively correlated with Factors I and II and probably representing testing atmosphere.

I.—Introduction.

Recently, Wallach and Kogan (1965) published an account of an intensive study of creative thinking in young children. This differed from most previous work in that the tests were administered in an individual rather than a group setting, considerable importance being attached to the presence of a "playful permissive task attitude" for the production of creative responses. A fundamental aspect of creativity itself was thought to be "The ability to give birth to associative content that is abundant and original yet relevant to the task rather than bizarre."

Before proceeding to more detailed analysis, the authors sought to establish the relative independence of the results of 'creativity' measures from those obtained from conventional tests of intelligence. For this purpose, fifteen tests were used. The first group consisted of five 'creativity' tests: 'Instances,' 'Alternate Uses,' 'Similarities,' 'Pattern Meanings,' and 'Line Meanings.' However, each of these tests was scored for both uniqueness of response and number of responses, making ten measures in all. A second group of ten 'g' tests comprised three WISC subtests: , 'Vocabulary,' 'Picture Arrangement,' 'Block Designs' and seven tests usually administered as group tests: the 'Sequential Tests of Educational Progress' (STEP)—five tests, and the 'School and College Ability' tests (SCAT)—two tests.

All testing was carried out within the same school, a group of 151 children (n girls=81 and n boys=70) of similar ' middle class ' background being used. From the results Wallach and Kogan derived an impressive series of intercorrelations, demonstrating a clear tendency for the ' creativity ' and ' g ' tests to form independent groups with little overlap at first sight. However, factorial data which would aid the further evaluation of particular tests was not supplied.

It was, therefore, decided to carry out an oblique factorization of the intercorrelations using the Promax method of Hendrickson and White (1964). This method involves powering the matrix of normalised Varimax loadings to obtain a criterion matrix of optimal simple structure. The original Varimax factors are then rotated to produce the best fit to this ' ideal ' matrix. The procedure generates a factor pattern and a factor structure together with the correlations between the oblique factors ; a second order oblique analysis of the latter is then usually made.

II.—Results.

A Principal Components Analysis of the intercorrelations was carried out and the four resultant significant factors rotated to Varimax criterion. After Varimax rotation, the percentage variances of these factors were 28·7, 23·8, 6·2 and 7·0, respectively (Total Variance 66·3). A full Promax analysis, including second order analysis, was carried out, but in the interests of brevity only the Factor Structure and a statement of the correlations between factors are given below :

TABLE 1

Factor Structure from Oblique Factorisation of Twenty Cognitive Tests (Wallach and Kogan, 1965)

	Test	Factor I	Factor II	Factor III	Factor IV
1.	Instances—uniqueness	—0·056	+0·509	—0·472	+0·312
2.	Instances—number	+0·110	+0·472	+0·487	+0·439
3.	Alternate Uses—uniqueness	+0·064	+0·836	+0·004	+0·218
4.	Alternate Uses—number	+0·175	+0·796	+0·283	+0·200
5.	Similarities—uniqueness	+0·033	+0·772	—0·106	+0·177
6.	Similarities—number	+0·179	+0·869	+0·130	+0·257
7.	Pattern meanings—uniqueness	+0·148	+0·653	+0·200	+0·099
8.	Pattern meanings—number	—0·033	+0·418	+0·661	+0·097
9.	Line meanings—uniqueness	+0·220	+0·717	+0·094	+0·354
10.	Line Meanings—number	+0·094	+0·742	+0·418	+0·164
11.	WISC vocabulary	+0·620	+0·192	—0·375	+0·494
12.	WISC picture arrangement	+0·182	+0·164	+0·001	+0·749
13.	WISC block design	+0·443	+0·037	—0·101	+0·631
14.	SCAT—verbal	+0·886	+0·165	—0·082	+0·282
15.	SCAT—quantitative	+0·860	+0·098	+0·097	+0·203
16.	STEP—mathematics	+0·831	+0·181	—0·182	+0·439
17.	STEP—science	+0·846	+0·167	—0·021	+0·417
18.	STEP—social studies	+0·887	+0·184	—0·099	+0·422
19.	STEP—reading	+0·859	+0·108	+0·026	+0·176
20.	STEP—writing	+0·889	+0·116	+0·111	+0·172

TABLE 2

Intercorrelations of Table 1 Factors.

I.	1·000	0·143	—0·031	—·0345
II.	0·143	1·000	0·116	—0·323
III.	—0·031	0·116	1·000	0·137
IV.	—0·345	—0·323	0·137	1·000

III.—Discussion.

Table 1 loadings represent the correlations of the tests with the oblique reference factors I–IV.

Factor I appears to be mainly a school attainment factor, tests 14–20 having remarkably similar loadings. As might be expected, WISC ' Vocabulary ' and ' Block Design ' load on this factor but none of the creativity measures possesses a substantial loading.

Factor II can be identified with creativity, all ten creativity measures having moderate to high loadings. The ' g ' tests all load positively on this factor, presumably reflecting the verbal nature of the tests and as may be seen in Table 1, Factors I and II are positively correlated.

Factor III is poorly defined but probably reflects scoring for number of responses as opposed to uniqueness or accuracy of response.

Factor IV is, however, the most interesting of the four. All twenty measures have positive loadings, with the three well established ' g ' tests, i.e., the WISC sub-tests, loading highest. There are thus some grounds for supposing that Factor IV may be a weak general factor. It is, however, negatively correlated with Factors I and II and is much more likely to represent the social atmosphere already described in contrast to the ' task ' requirements of the situation.

IV.—Conclusions.

The above analysis has proved to be extremely useful for the construction of ' creativity ' items for intelligence tests. It was not originally intended to prove or disprove the inferences drawn by Wallach and Kogan, although the results tend to support their choice of procedure. However, the multifactorial nature of ' creativity' data is once more demonstrated despite the presence of two apparently near orthogonal and easily identifiable sets of measures.

Acknowledgement.—The author wishes to acknowledge the help of Professor F. W. Warburton for his useful criticisms and advice on the preparation of this paper.

V.—References.

Hendrickson, A. E., and White, P. O. (1964). Promax . a quick method for rotation to oblique simple structure. *Brit. J. Stat. Psychol*, 17-1, 65-70.

Wallach, M. A., and Kogan, N. (1965). *Modes of Thinking in Young Children : A Study of the Creativity-Intelligence Distinction.* New York . Holt, Rinehart and Winston.

(*Manuscript received 3rd November, 1966*)

THE RELATIVE VALUE OF TWO HIGH LEVEL INTELLIGENCE TESTS, ADVANCED LEVEL, AND FIRST YEAR UNIVERSITY EXAMINATION MARKS FOR PREDICTING DEGREE CLASSIFICATION

By G. W. PILKINGTON and G. J. HARRISON
(*Department of Psychology, University of Sheffield*)

I.—Introduction.

Many workers in this country have enquired into the predictive value either of intelligence test scores or of marks or grades in school-leaving examinations for university students' later academic achievements, usually as measured by degree class, if any, subsequently attained. Examples of such investigations are papers by Williams (1950), Himmelweit and Summerfield (1951), Gould and McComisky (1958), Richards and Wilson (1960-61), Austwick (1960-61), Petch (1961), Furneaux (1961), Himmelweit (1963), and Nisbet and Welsh (1965-66). Critical reviews of the literature have been made by Eysenck (1947) and Himmelweit (1950) on mainly American findings, and by Warburton (1952), Dale (1954), Montgomery (1962) and Kelsall (1963), with more reference to British work. With some qualifications, one conclusion that may be drawn from these investigations is that marks or grades in school-leaving examinations—in particular 'A' level grades—provide a useful, though far from perfect, indication of a candidate's likely academic achievement. The results of work on intelligence tests are, however, less conclusive, in spite of the early optimism of Eysenck (1947) and of the challenging results reported by Himmelweit and Summerfield (1951) and Himmelweit (1963). One reason for this, suggested by Himmelweit (1950), Valentine (1961) and Montgomery (1962), is that the relatively narrow range of ability of applicants to university in this country makes intelligence tests standardised on the general population of little use as predictors of candidates' probable academic achievements. But an obvious possiblity is that tests designed to measure high level intelligence might prove to be better predictors. Such tests have been reported by Heim (1947), Foulds and Raven (1950) and Valentine (1954), and some evidence that they might be of value as selection devices has been provided by Wallace (1952), Heim and Watts (1960), Valentine (1961) and Kelvin, Lucas and Ojha (1965) But, so far, no published accounts of efforts to compare directly the relative value of high level intelligence tests and of 'A' level grades as predictors of students' later academic performance have appeared, although comprehensive findings of this kind will, of course, ultimately result from Furneaux's Nuffield investigation (see Furneaux, 1961). In the meantime, the study reported below, which attempts to make such a comparison, may be of some interest.

II.—METHOD OF INVESTIGATION AND STATEMENT OF RESULTS.

Between the years 1955 and 1963, inclusive, 252 students took the first year course in Psychology at the University of Sheffield. These students then went on to read for degree courses in a wide variety of subjects within the Faculties of Arts and of Social Sciences. All of these students were given the A.H.5 group test of high-grade intelligence and 153 the Valentine reasoning tests for high levels of intelligence; 243 of the students had been admitted to the University on the strength of 'A' level marks or grades in various school subjects, and for all such students a mean 'A' level mark was calculated. The statistical objections to averaging such marks or grades are obvious (although see Anderson, 1961), and the principal reason for adopting such a dubious device was simply that it is commonly used by university admissions officers as a means for selecting candidates. In all cases, this mean was based on the marks awarded to a candidate in his first attempt at two or more 'A' level subjects, for, as Dale (1954) (1961) and others have emphasized, such marks have more predictive value than those gained at later attempts. 'A' level grades, whether numerical or alphabetical, were translated into marks according to a conversion scale used by the university admissions officer, thus: Grade A equals 75 marks; B equals 65; C equals 57; D equals 52; E equals 45; O and F equal 35; and similarly for numerical grades. Clearly, such a conversion is highly questionable, and the only justification for our adopting this procedure here is that it is in fact one used by selectors. In due course, marks became available for the students from examinations in three or four subjects taken at the end of their first year at the university. This examination will henceforth be referred to as the first university examination. For all students who took this examination (three graduates taking a further first degree were exempted from it) a mean first university examination mark was worked out. Again, only marks gained at a first attempt at the examination were used, and, as before, the main reason for averaging was that such a mean mark is, in fact, worked out from the unstandardised marks in the different subjects and then used by the university examiners as a means of selecting students for different courses. By 1966, 215 of the students in the group had been awarded degrees at their first attempt, and thirty-one had failed or been sent down for academic reasons (giving a failure rate of 12·6 per cent.). The remaining six had either left for non-academic reasons without taking their degree examinations or had yet to take their degrees. The 246 students for whom degree results were available were then divided into four categories, according to degree classification attained : (i) those who had gained first or upper-second class honours in honours degree examinations and those who were awarded first or second class honours in general degree examinations ; (ii) those who had gained lower-second class honours in honours degree examinations and those who were awarded a pass division one in general degree examinations ; (iii) those who gained third class honours or a pass in honours degree examinations and those who were awarded a pass division two in general degree examinations ; (iv) those who had failed honours and general degree examinations and those who were were sent down for academic reasons without being allowed to complete their degrees.

The students' scores on the two tests and their mean marks in the three examinations, together with the date of their year of entry to the university, were then subjected to a matrix correlation analysis using the technique of product-moment correlation.* The results are given in Table 1 below.

TABLE 1

MATRIX OF CORRELATIONS BETWEEN HIGH LEVEL INTELLIGENCE TEST SCORES AND MEAN MARKS OR CLASS IN SCHOOL AND UNIVERSITY EXAMINATIONS.

Variable	Mean 'A' level mark	Valentine score	A.H.5 score	Mean first University examina-mark	Degree Class
Year of entry	+0·376***	+0·116	+0·154*	+0·171*	+0·113
d.f.	241	151	250	247	244
Mean 'A' level mark ..		+0·178*	+0·118	+0·241*	+0·300**
d.f.	..	145	241	240	235
Valentine score			+0·473***	+0·115	+0·276**
d.f.			151	149	151
A.H.5 score				+0·114	+0·181**
d.f.				247	244
Mean first University examination mark ...					+0·524***
d.f.					241

Significance levels : * $=p<0\cdot05$. ** $=p<0\cdot01$. *** $=p<0\cdot001$.

Multiple regression analyses were made using all possible combinations of test scores and of examination marks in predicting degree class. These analyses employed the data from the 146 students who had scores on all five variables. The simple correlations of each of the four predictor variables with degree class for this sample of students were also worked out for comparison purposes The results are shown in Table 2.

---

* In an earlier version of this research, forty-one variables were intercorrelated. The extra variables comprised 'A' level, first year university and degree examination marks in individual school and university subjects. But numbers in most cases were so small that the results, although suggestive and of some interest, were rarely significant. The findings are, however, reported and discussed in full by Stafford (1965).

TABLE 2

REGRESSION ANALYSES USING FOUR PREDICTORS OF FINAL DEGREE CLASSIFICATION. (N=146)

Predictor variables	Correlation coefficients R	Values of beta coefficients, multiplied by (10^{-4}), for each predictor in multiple regression analyses: Mean 'A' level mark 1	Valentine score 2	A.H.5 score score 3	Mean mark in first University examination 4
1 2 3 4	0·5691	3099	930	–1727	5211
1 2 4	0·5596	2997	1081		5218
1 3 4	0·5377	3365		–401	5418
2 3 4	0·5262		1039	–1479	6122
1 2 3	0·4335	4668	1075	–1753	
1 4	0·5372	3318			5406
2 4	0·5186		843		6101
3 4	0·4830			42	6444
1 2	0·4200	4566	848		
1 3	0·3755	5050		–211	
2 3	0·2680		1296	–135	
	r				
4	+0·4828	p<0.001			
1	+0·3752	p<0·001	All multiple correlations are		
2	+0·2526	p<0·01	significant at beyond the		
3	+0·0392	n.s.	1 per cent. level.		

Finally, a method of analysis of the findings, somewhat similar to those adopted by Valentine (1954) (1961) and Petch (1961), was attempted. The mean scores of the students on the two tests and the two examinations were worked out according to class of degree attained, and then were subjected to analysis of variance. Table 3 shows the means, *F* ratios and levels of significance.

TABLE 3

MEAN SCORES AND MARKS GAINED BY STUDENTS ON TWO TESTS AND TWO EXAMINATIONS IN RELATION TO DEGREE CLASSIFICATION.

	Degree Class (i)	Degree Class (ii)	Degree Class (iii)	Degree Class (iv)	F ratios	Signi-ficance level
Mean first University examination mark . . .	54·42	50·75	48·19	40·71	32·96	<0·001
	N=49	N=107	N=57	N=30		
Mean 'A' level mark . . .	57·87	55·09	52·37	50·32	7·79	<0·001
	N=47	N=104	N=57	N=29		
Valentine score	42·88	36·43	32·60	30·05	4·36	<0·01
	N=25	N=67	N=41	N=20		
A.H.5 score	39·66	35·51	34·59	35·29	4·99	<0·01
	N=51	N=107	N=57	N=31		

All statistical analyses reported were carried out by computer.

III.—DISCUSSION OF RESULTS.

Data concerning the students' year of entry to the university were included in the correlation analysis shown in Table 1 in an attempt to test the common supposition that the 'quality' of students at the university has risen in recent years because of the increased competition for university places. Some support is given to this hypothesis, in that significant positive correlations emerge between year of entry and three of the five variables. Especially noteworthy is the relatively high correlation between year of entry and mean 'A' level mark. This is consistent with the poilicy the university has adopted to cope with the increasing numbers of applicants-that of gradually raising over the years the minimum 'A' level admission requirements to these two Faculties. It is also of some interest that there is a significant, if lower, positive correlation between year of entry and mean first university examination mark This suggests that the university examiners in these two Faculties have not been entirely insensitive to the rising academic calibre of their students.

The remaining figures in Table 1 do little to support the idea that high-level intelligence tests may be more useful as selection devices than the conventionally used 'A' level grades, since neither test correlates as highly with degree class as does the mean 'A' level mark. However, the difference between the correlation of ·300 for the mean 'A' level mark and degree class and that of ·276 for Valentine score and degree class is not statistically significant, and it should be borne in mind that the correlation between 'A' level mean mark and degree class may itself be spuriously high by reason of the fact that most students score within the middle range on both variables. But perhaps the most striking features of Table 1 are the relatively high correlations between mean first university examination mark and degree class and the surprisingly low correlation between A.H 5 score and degree class. The latter compares poorly with other similar investigations of this test, for example those by Heim and Watts (1960), Banks (quoted by Valentine, 1961) and Kelvin, Lucas and Ojha (1965). Part of the explanation for this may be that the A H.5, in our sample at least, seems to predict degree classification negatively rather than positively : low scorers on this test (that is, those who score one standard deviation or more below the mean of the sample) very rarely achieve the top degree classification, whereas high scorers on this test (that is, those who score one standard deviation or more above the mean of the sample) are almost equally as likely to fail their degree as are low scorers. In other words, the A.H.5 may be more useful negatively as a device for excluding candidates unlikely to attain a good class degree than positively as a means of predicting high academic achievement.*

The data presented in Table 3, although resulting from an entirely different kind of analysis from that used for the findings shown in Table 1, nevertheless, lead to broadly similar conclusions. Again, the mean first university examination mark discriminates by far the most effectively between the different degree classes and again this is followed by the mean 'A' level mark. In comparison, the two tests, perform relatively poorly, although the Valentine is handicapped by being based on rather small numbers. It is at least encouraging that the mean scores on this test for the different degree classes are somewhat similar (in so far as they can be compared) to those originally reported by Valentine himself (1954) (1961). It is noteworthy that, in contrast, the A.H.5 discriminates hardly at all between the bottom three degree classes.

Finally, the figures presented in Table 2 appear to be compatible with those in the other two tables. Combinations which include the mean 'A' level mark and especially the mean first university examination mark tend to produce the higher multiple correlations with degree class, whereas combinations which include the Valentine and especially the A.H.5 tend to produce the lower multiple correlations.

---

*A detailed analysis of the relationship between high and low scores (as defined above) on the two tests and the two examinations and degree classification is not presented here. But, with the exception of the A.H.5, which discriminates in the asymmetrical way indicated, the results of this analysis are closely comparable with the findings presented in the three tables shown above.

Similar conclusions are also suggested from inspection of the beta coefficients. The simple correlations reported are similar to those found for the whole sample in Table 1, except that the formerly low but significant correlation for the A.H.5 now disappears. But the most interesting feature of Table 2 is that a combination of all four predictor variables results in a correlation of only 0·5691, scarcely better than the simple correlation yielded by the mean first university examination mark. In other words, the two measures of academic achievement and the two measures of intellectual potential, even when taken together, account for less than a third of the variance in degree class obtained.

IV.—CONCLUSIONS.

Clearly, the findings reported must be interpreted with considerable caution. They are based on a relatively small sample of students drawn from only two faculties of one particular university, and there may also be an additional sample bias—all the students opted to read psychology for at least one year ! Moreover, the criterion of degree class awarded, against which the test scores and examination marks are matched, is notoriously unreliable, as has been amply documented by Dale (1954) (1958-59). In addition, the research was necessarily confined to candidates actually admitted to the university. No attempt has been made to allow for these last two factors by correcting for attentuation and range (see Furneaux, 1961, p. 91 *seq*.).

Granted these limitations, however,the findings of the present enquiry suggest the following tentative conclusions. First, the two high level intelligence tests predict degree classification no better than the conventionally used 'A' level marks ; indeed, the A.H.5 predicts it significantly less well for the smaller sample considered in Table 2. Even combining the two test scores with mean 'A' level mark results in only a slight improvement on the correlation coefficient found for the latter measure alone. Use of these two tests, particularly of the A.H.5, for selection purposes cannot, therefore, be recommended on the basis of these results (except that the A.H.5 may have some negative predictive value, as suggested earlier). On the other hand, the students' motivation to perform well was likely to have been much higher when taking their 'A' level and university examinations than when taking the two tests (on which they were told nothing serious would depend), and it would be interesting to discover whether the above results would be replicated if the two tests were given as part of a university entrance selection procedure.
Second, the findings reported confirm that marks in school-leaving examinations provide a useful, if imperfect, indication of a candidate's academic quality ; this is in agreement with the results of many other workers, as mentioned earlier. Third, it is clear that marks awarded in the first university examination have a useful predictive value for final degree class, at least in the context of this particular investigation, and one which is significantly higher ($p<0{\cdot}05$) than that attained by the 'A' level marks. This is, perhaps, because this examination has more in common with the final examination, for example in terms of markers and their conventions and expectations, than have 'A' level examinations. Even so, it is doubtful whether this finding entirely justifies the common university policy of using ***unstandardised*** marks gained in examinations at the end of the first year as a basis for allocating students to different levels and kinds of courses (for example, Honours and General) and, in cases of bad failure, for excluding students from further attendance at the university. Finally, and possibly most important from the point of view of future research, none of the measures used correlated very highly with degree class awarded, either singly or in combination ; similarly none of them discriminated as clearly between the different degree classes as might have been desired. Other cognitive measures, either of a test or an examination variety, might conceivably have fared better. So might a scholastic aptitude test, as advocated by Oliver (1961-62) But until this can be shown, it seems plausible to suppose that a sizeable proportion of the variance in degree class attained may be attributable to variables which cannot be measured by cognitive tests or examinations. It is significant that, in recent years, much interest has been shown in the non-cognitive factors which may affect academic performance, especially personality and temperamental variables.

Such publications as those by Hopkins, Malleson and Sarnoff (1958), Lynn (1959), Lynn and Gordon (1961), Savage (1962), Furneaux (1962-63), Malleson (1963), Dale (1963) and Hudson (1963), (1966), sufficiently demonstrate this trend of interest. Fully agreed findings have yet to emerge from this difficult and controversial field of research (c.f. Biggs, 1962), but until they do, it may be supposed that university selectors will continue to work, if not in the dark, then certainly in the twilight.

AKNOWLEDGEMENTS.—The authors would like to thank Miss M. B. Stafford, who carried out an earlier version of this research, and Miss V. G. Haycock and Mr. T. Green, who provided much appreciated help in the computer analysis of the data.

V.—REFERENCES.

ANDERSON, N. H. (1965). Scales and statistics : parametric and nonparametric. *Psychol. Bull.*, 58, 305-316.

AUSTWICK, K. (1960-61)' G.C.E. to B.A. *Univ. Quart.*, 15, 64-71.

BIGGS, J. B. (1962). The relation of neuroticism and extraversion to intelligence and educational achievement. *Brit. J. Educ. Psychol.*, 32, 188-195.

DALE, R. R. (1954). *From School to University*. London : Routledge and Kegan Paul.

DALE, R. R. (1958-59). University standards. *Univ. Quart.*, 13, 186-195.

DALE, R. R. (1961). Critical notice of W. D. Furneaux's *The Chosen Few*. *Brit. J. Educ. Psychol.*, 31, 303-305.

DALE, R. R. (1963). Reflections on the influence of social class on student performance at the University. *Sociol Rev.*, Monograph No. 7, 131-140.

EYSENCK, H. J. (1947). Student selections by means of psychological tests—a critical survey. *Brit. J. Educ. Psychol.*, 17, 20-39.

FOULDS, G. A., and RAVEN, T. C. (1950). An experimental survey with progressive matrices (1947). *Brit. J. Educ. Psychol.*, 20, 104-110.

FURNEAUX, W. D. (1961). *The Chosen Few*. London : Oxford University Press.

FURNEAUX, W. D. (1962-63). The psychologist and the university. *Univ. Quart.*, 17, 33-47.

GOULD, E. M., and MCCOMISKEY, J. G. (1958). Attainment level on leaving certificate and academic performance at university. *Brit. J. Educ. Psychol.*, 28, 129-134.

HEIM, A. W. (1947). An attempt to test high-grade intelligence. *Brit. J. Psychol.*, 37, 70-81.

HEIM, A. W., and WATTS, K. P. (1960). An enquiry into the relationship between university career and scores on a high grade intelligence test. *Cambridge Review*, 81, 463-469.

HIMMELWEIT, H. T. (1950). Student selection—an experimental investigation : I. *Brit. J. Sociol.*, 1, 328-346.

HIMMELWEIT, H. T. (1963). Student selection—implications derived from two student selection enquiries. *Sociol Rev.*, Monograph No. 7, 79-98.

HIMMELWEIT, H T , and SUMMERFIELD, A. (1951). Student selection—an experimental invesgitation : II and III. *Brit. J. Sociol.*, 2, 59-75, 340-353.

HOPKINS, J., MALLESON, N., and SARNOFF, I. (1958). Some non-intellectual correlates of success and failure among university students *Brit. J. Educ. Psychol.*, 28, 25-36.

HUDSON, L. (1963). The relation of psychological test scores to academic bias. *Brit. J. Educ. Psychol.*, 33, 120-131.

HUDSON, L. (1966). *Contrary Imaginations*. London : Methuen.

KELSALL, R. K. (1963). University student selection in relation to subsequent academic performance : a critical appraisal of the British evidence. *Sociol Rev.*, Monograph No. 7, 99-115.

KELVIN, R. P., LUCAS, C. P., and OJHA, A. B. (1965). The relation between personality, mental health and academic performance in university students. *Brit J. Soc. Clin. Psychol.*, 4, 244-253.

LYNN, R. (1959). Two personality characteristics related to academic achievement. *Brit. J. Educ. Psychol.*, 29, 213-216.

LYNN, R., and GORDON, I. E. (1961). The relation of neuroticism and extraversion to intelligence and educational attainment. *Brit. J. Educ. Psychol.*, 31, 194-203.

MALLESON, N. (1963). The influence of emotional factors on achievement in university education. *Sociol Rev.*, Monograph No. 7, 141-159.

MONTGOMERY, G. W. G. (1962). *Student Achievement and Selection to University*. University of Edinburgh : Applied Psychology Unit Report.

NISBET, J., and WELSH, J. (1965-66). Predicting student performance. *Univ. Quart.*, 20, 468-480.

OLIVER, R. A. C. (1961-62). The selection of university students : a ' scholastic aptitude test ' ? *Univ. Quart.*, 16, 264-273.
PETCH, J. A. (1961). *G.C.E. and Degree, Part* 1. Manchester : Joint Matriculation Board.
RICHARDS, J. P. G., and WILSON, A. J. C. (1960-61). A-level and pass degree in physics. *Univ. Quart.*, 15, 389-392.
SAVAGE, R. D. (1962). Personality factors and academic performance. *Brit. J. Educ. Psychol.*, 32, 251-253.
STAFFORD, M. B. (1965). *A Comparison of Different Methods of Predicting Students' Achievement in University Examinations.* University of Sheffield ; B.A. degree dissertation.
VALENTINE, C. W. (1954). *Handbook of Instructions for Reasoning Tests for Higher Levels of Intelligence.* Edinburgh : Oliver and Boyd.
VALENTINE, C. W. (1961). The use of a new reasoning test for selection of university and training college students. *Brit. J. Educ. Psychol.*, 31, 227-231.
WALLACE, J. (1952). Results of a test of high-grade intelligence applied to a university population. *Brit. J. Psychol.*, 43, 61-69.
WARBURTON, F. W. (1952). *The Selection of University Students.* Manchester : Manchester University Press.
WILLIAMS, E. R. (1950). An investigation of the value of higher school certificate results in predicting performance in first-year university examinations. *Brit. J. Educ. Psychol.*, 20, 83-98.

(*Manuscript received 7th November*, 1966)

THE PERFORMANCE OF AMERICAN PUPILS ON THE COLOURED PROGRESSIVE MATRICES

BY ZITA M. CANTWELL
(*Brooklyn College of the City University of New York*)

This study was undertaken as a normative descriptive investigation of the performance of students from two selected public elementary schools, grades 2 through 5, on the Coloured Progressive Matrices, revised order, 1956 (Raven, 1963).

I.—PLAN.

The Sample. The schools from which the students were drawn were differentiated by (1) the cultural and economic styles of the neighbourhoods they served, (2) the ethnic group membership of the two student bodies, and (3) the language development of the students.

The students in School I come from a stable neighbourhood of reasonable economic prosperity. Family expectations for the students are, generally speaking, high. The students are verbal ; language is used to express ideas, emotions, solutions to problems. The ethnic group membership of the population in this school was approximately : Other—94 per cent.; American Negro—3 per cent. ; Puerto Rican—3 per cent. (Ethnic group membership data were obtained from the New York City Board of Education school census, which uses these three categories.)

School II is located in a rapidly changing neighbourhood. Family expectations, in terms of student's completing school, are relatively high ; but the families themselves are not always able completely to implement such expectations. The students are not highly verbal ; they do not grow in a verbal environment; gesture, movement or action accompanied by a minimum of speech are used to transmit ideas, emotions, problem solutions. The ethnic group membership of the population in this school was approximately : Other—3 per cent. ; American Negro—94 per cent ; Puerto Rican—3 per cent.

The total sample included 727 students (358 boys, 369 girls) from School I and 807 students (417 boys, 390 girls) from School II.

Procedure. In October and November of the school year, the Coloured Progressive Matrices was administered as an untimed group test to the students in Schools I and II. Descriptive data, including ethnic group membership distributions in the two schools, were obtained from school records.

II.—STATEMENT OF RESULTS.

Reliability. The reliability of the Coloured Progressive Matrices for the two samples was estimated by applying Kuder-Richardson formula 20. The Matrices appears equally reliable for both school samples : School I : $r_{tt} = {\cdot}837$ (N=727) ; School II : $r_{tt} = {\cdot}822$ (N=807).

Performance on Coloured Progressive Matrices, Schools I and II.

(1) Data from both schools yielded similar Coloured Progressive Matrices total score distribution patterns, but the mean total scores between the two schools by grade, were significantly different. The performance level in School I was higher that that in School II (see Table 1).

TABLE 1

COLOURED PROGRESSIVE MATRICES MEAN TOTAL SCORES AND STANDARD DEVIATIONS BY SCHOOL AND GRADE ; DIFFERENCES BETWEEN SCHOOL MEANS BY GRADE AND DIFFERENCES BETWEEN GRADE MEANS WITHIN EACH SCHOOL.

Grade	School I Mean	Diffs. between Grade Means	s d.	N	School II Mean	Diffs. between Grade Means	s.d.	N	Means I–II
5	28·28		5·69	192	22·10		5·79	174	6·18*
		1·43*				3·19*			
4	26·85		5·39	179	18·91		5·48	169	7·94*
		3·18*				0·65			
3	23·67		5·41	183	18·26		5·25	248	5·41*
		2·61*				1·38*			
2	21·06		6·03	173	16·88		4·44	216	4·18*

*$p < {\cdot}01$

(2) Within both School I and School II, the Coloured Progressive Matrices mean total scores obtained by the boys in each grade were consistently higher than mean total scores obtained by the girls

Item Analysis. Analysis of the scores obtained by the total sample indicated that the arrangement of questions in the sub-tests of the Coloured Progressive Matrices does not entirely implement a cyclical pattern of test design. The test may contain too many simple items, items of misplaced difficulty level, and abrupt increases in item difficulty levels within the subtests.

III.—REFERENCE.

RAVEN, J. C. (1963). *Guide to Using the Coloured Progressive Matrices (revised order, 1956).* London : H. K. Lewis.

(*Manuscript received 8th June,* 1967)

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES

AN INVESTIGATION INTO THE READING EFFICIENCY OF STUDENTS AT A TECHNICAL TEACHER TRAINING COLLEGE

By RALPH LIONEL JAMES

(*Summary of a Dissertation submitted in part fulfilment of the requirements for the M.A. (Education) Degree of the University of London*, 1965)

One of the objectives of this study was to discover the value of the use of the Harvard University Reading Course[1] with students in a technical teacher training college. The students are mainly in their late 20's, 30's or 40's, with a few over 50 years of age. Their academic qualifications on admission to the college include City and Guilds Certificates, Short-hand and Typewriting Certificates and National Certificates in Engineering, as well as University Degrees and other qualifications. They vary considerably also in Intelligence Test scores on entry.

The training consisted of eight sessions, during which eight films and eight reading passages from the Harvard course, together with their related 'multiple-choice' comprehension tests, were administered to the members of the Experimental Groups. The films used were No. 1, followed in successive weeks by Nos. 6–13, omitting No. 12. The sessions were at weekly intervals, except for a three-week break in the middle of the training, while the students were on teaching practice.

Apart from the films and the reading passages, the training also included the enumeration of the 'skills of efficient reading' during the initial diagnostic testing session and their occasional repetition during the course of instruction. These skills were enumerated as: No sub-vocalisation (although conscious attention to avoiding this may be questioned as a method of learning); avoidance of regressive eye movements; avoidance of head movements; wide eye-span units; reading ideas rather than words; different speeds for different purposes. The last of the skills was stressed paricularly as being indicative of the good reader.

The results of the investigation are summarised in Table 1:

In general, these results confirm what has generally been found in other investigations (e.g., Poulton[2]). It is evidently possible, even by a very abbreviated course of training of this kind to raise the reading speed of trainees substantially with no significant change in comprehension scores compared with a control group not undergoing the training. From the table it can be seen that the significant gains in speed occurred in the groups which were initially below the college average in 'reading efficiency' with respect to speed, comprehension or both. It is argued in the Dissertation that the multiple-choice method of testing remembering had a most important bearing on this result (McLaughlin, G. H.[3]).

In interpreting the results, it should be borne in mind that the reliability coefficient for the speed scores of the Control Group (tested on a counter-balanced order basis) was +0·67 (S.E. = ·06). Allowing for the factors tending to reduce this reliability, such as 'change of set,' this degree of reliability makes the test of diagnostic significance. The comprehension scores for the Control Group also had a high degree of reliability—variations of more than ±·2 being rare among the students. This degree of variability may well result from 'chance' factors on this type of four-alternative multiple-choice test. A high degree of consistency in comprehension test scores was also characteristic of the Experimental Groups throughout the training.

TABLE 1

SUMMARY OF RESULTS OF APPLYING READING PROGRAMME.

	Number in group	Beginning of term test—A		End of term test—B		
		Mean speed w.p.m.	Mean Comp. Score out of 10	Mean Speed w.p.m.	Mean Comp. Score out of 10	% gain in speed
Experimental Groups : Science Group	15	171·52	6·87	206·84	6·67	20·6
Secretarial Group 3	13 (2 missed Test A or B)	181·22	5·64	203·82	6·27	12·5
Catering Group..	13	135·31	5·92	260·18	5·92	92·3
Printing Group..	11 (2 missed Test B)	170·84	5·11	223·57	5·78	30·9
Clothing and Allied Subjects Group........	10 (1 missed Test A)	137·45	5·44	222·14	6·11	61·6
Control Group : ' Rest of College '	189	168·81	6·16	184·83	6·59	9·5

No change in mean comprehension score of the experimental groups was statistically significant. Of the gains in speed only the last three groups (the Catering, the Printing and the Clothing Groups) were significant at the ·05 level.

It is this consistency in comprehension test scores in both the Control and Experimental Groups which gives the investigation its principal significance. It is suggested that consistently low comprehension scores on reading material of the order of difficulty of the Harvard tests and course material indicate diagnostically a low probability of successful study reading of the type of texts involved in the study of the Principles of Education, which was part of the curriculum of the college at the time of the study.

To meet the possible criticism that consistently low comprehension scores may be the outcome of an unnatural reading set induced by a ' Speed Reading Course,' it is necessary to look at both Speed and Comprehension Test scores in order to verify that on the diagnostic testing low comprehension test scores were not ' offset ' by high speed scores. The relevant concept is that of reading efficiency, embracing both speed and comprehension, rather than speed or comprehension separately.

On the diagnostic testing 100 students out of a college population of 278 were found to have mean speed scores of 144 words per minute or less *or* mean comprehension scores of less than 50 per cent. In case these standards might be considered too high as a basis for diagnosing an ' unsatisfactory scores ' group with a low probability of successful study reading, a check was made to ascertain how many of these students would have been omitted from this group if 'very satisfactory' comprehension scores had been allowed to permit exclusion regardless of speed scores and if students with one 'satisfactory' combined speed/comprehension score were also excluded. Ten students were found to come in the first of these categories and five in the second. If we were to deduct these students from the 100 members of the 'unsatisfactory scores' group, we are still left with eighty-five members of this category out of a college population of 278.

It is concluded on what appear to be very lenient criteria, therefore, that roughly one-third of the student population have a low probability of successful study reading of texts on education. Even allowing for the limitations of the method of diagnostic assessment, this would appear to have quite considerable implications for curriculum design.

In relation to the diagnostic use of the Harvard Tests, it should be noted that the three Experimental Groups which were originally chosen randomly from all groups in the college, were given the A H.4 Intelligence Test. Their results correlated badly with the Harvard tests. It appears, therefore, that the Harvard tests were measuring something distinct from this well-known test, and something which appears manifestly to be a closer simulation of study skills

REFERENCES.

[1] PERRY, W. G., Jnr., and C. P. WHITLOCK, (1949). *Instructor's Manual.* Harvard University Reading Course, Harvard University Press.

[2] POULTON, E. C. (1961). British courses for adults on effective reading. *Brit. J. Educ. Psychol.*, 31, 128-137.

[3] McLAUGHLIN, G. H. (1964). The potential for reading faster. *New Education*, 1, 35-37.

(*Manuscript received 7th December*, 1966)

THE VALIDITY OF OBJECTIVE TESTS IN ENGLISH

BY F. H. SPARROW

(*Summary of thesis submitted for the degree of M.A. (Ed.), University of Southampton,* 1966)

I.—AIM.

To determine the validity of objective tests in English in the context of an investigation into the overlap area between the Certificate of Secondary Education and the Ordinary level of the General Certificate of Education.

II.—THE SAMPLING.

For preliminary trials of the types of test, samples of about 100 G.C.E. and C.S.E. candidates were taken in the Southampton area. For the main experiment, 889 children from two G.C.E. boards, and 762 from two C S.E. boards were used.

III.—METHOD.

Preliminary trials were used to determine which types of test showed the highest agreement with the criteria of teachers' assessments and examination results and to refine the test items in the light of item analysis. From this procedure, a test battery consisting of four types of test was produced, with a time-allowance of 50 minutes. A standardised vocabulary test with a time-allowance of 10 minutes, a 45-minute test of continuous writing, and a 60-minute test of general ability were added to the battery.

IV.—Types of test.

Comprehension. A multiple-choice test of twenty items, based on an extract from a modern author. Some questions required meanings of phrases, others required a sympathetic appreciation of the attitude of the author.

Interlinear. A correction exercise of twenty items, based upon those aspects of grammar, syntax and usage which were within the experience of the fifth-year secondary school pupil.

Critical Judgment. A multiple-choice of ten items requiring the scrutiny and critical evaluation of advertiser's prose.

Paragraph Organisation. Two groups of 'scrambled' sentences requiring re-arrangement in the best order. (Fifteen items in all.)

Vocabulary. The Watts /Vernon vocabulary test of thirty-five items.

V.—Results.

With the exception of the vocabulary test, each of the types of test included in the final battery was used in at least one preliminary trial. The validity coefficient of each battery using teacher's assessments as criteria was found to be ·481, ·532, ·618 and ·650. Using an experimental examination in English (C.S.E.) and O-level English Language as criteria, the values were ·515 (N=75)and ·594 (N=25), respectively. Before revision of items in the light of item analysis, each type of test showed positive agreement with teachers' assessments, ranging from ·312 (for the highly tentative test of critical judgment) to ·689 (for the operationally-successful interlinear test).

The agreement between the tests, the essay marks and the scores on the test of general ability, and the examination grades, is shown in Table 1.

TABLE 1

Agreement between Tests and Examination Grades.

	G.C.E. Boards		C.S.E. Boards	
	A N=625	B N=264	X N=554	Y N=208
Comprehension	·442	·305	·324	·254
Interlinear	·510	·449	·490	·246
Critical Judgment	·146	·265	·148	·254
Paragraphs	·293	·167	·241	·177
Vocabulary	·469	·393	·416	·278
Objective test total score	·597	·478	·521	·382
Essay	·373	·368	·443	·280
General Ability	·411	·282	·293	·172

Two different methods of pooling the essay and objective test scores were investigated. In the first the critical judgment and paragraphs test scores were omitted, the vocabulary test scores scaled to a maximum of 20 and the essay marks doubled to a maximum of 20. In the second, the objective test selection and weighting were as for method 1, but the essay scores were weighted to a maximum of 40. The results obtained are set out in Table 2.

TABLE 2

	A	B	X	Y
Method 1	·624	·546	·587	·363
Method 2	·593	·527	·586	·349

When the tests were used as calibrating instruments, performances were expressed in the form of percentile levels in order to eliminate regression effects. For example, a score of 55 on the English objective test in Board A occurs at the bottom of Grade 6 (at a percentile level of 28·3) and in Board B the same score (55) on the same test also occurs at the bottom of Grade 6 (at a percentile level of 26·8). The equivalent scores at the bottom of Grade 1 in the C.S.E. Boards X and Y were 56 and 59. The overlap between the other grades in terms of the objective test scores was determined in the same way and the same procedure was adopted for the essay and the general ability test scores. The overlap scores for all grades are set out in Table 3.

TABLE 3

OVERLAP OF GRADES OF G.C.E. AND C.S.E.

G.C.E. Boards						C.S.E. Boards					
A			B			X			Y		
Rank %	Bottom of Grade	Objective Test Score	Rank %	Bottom of Grade	Objective Test Score	Rank %	Bottom of Grade	Objective Test Score	Rank %	Bottom of Grade	Objective Test Score
89·1	1	73	94·6	1	77						
83·0	2	70	86·3	2	72						
59·5	3	63	67·4	3	65						
49·2	4	60	53·7	4	62						
40·3	5	58	38·2	5	58						
28·3	6	55	26·8	6	55	85·7	1	56	93·7	1	59
16·3	7	51	14·3	7	49	59·7	2	47	78·3	2	51
7·0	8	47	6·4	8	45	32·1	3	39	58·1	3	46
						11·9	4	33	9·0	4	30
						3·2	u	26	1·9	u	25

NOTE.—The alignment of G.C.E. grades 7 and 8 with C.S.E. grades 2 and 3 is for convenience in tabulation and does not imply equivalence.

(*Manuscript received 3rd December*, 1966)

THE OPINIONS ON CO-EDUCATION OF A GROUP OF SECONDARY MODERN SCHOOL CHILDREN IN THE MANCHESTER AREA, THEIR TEACHERS AND THEIR PARENTS

By AUDREY NICHOLLS

(*Abstract of thesis presented for the degree of M.Ed., in the University of Manchester,* 1966)

I.—AIM.

The aim of the research was to obtain the opinions on co-education of a group of secondary modern pupils, their parents and their teachers, and to examine the relationships between the opinions offered. It was hypothesized that the opinions on co-education would be related to age, sex, socio-economic level and type of education received.

II.—THE SAMPLE.

Twelve schools in the Manchester area were used : four boys', four girls' and four mixed. The schools were matched on the basis of size, type of locality and the facilities available. Children were selected from the second, third and fourth years of each school by random sampling, 270 pupils, 132 girls and 138 boys, being selected; 197 of the parents of these children formed the parent sample (almost 73 per cent. of the parents) and 152 teachers formed the teacher sample (almost 71 per cent. of the teachers).

III.—THE QUESTIONNAIRES.

Three questionnaires were constructed, one for pupils, one for their parents and one for the teachers. All were answered anonymously. The children were asked to give the name of their school, their form, age, parents' occupation and number of siblings. They were asked whether they would choose a mixed or a single-sexed school if they went to live in another town and then to select from a given list the reasons which made them choose the type of school. The parents were asked to give their age and occupation, and to state whether they had attended a mixed or single-sexed school. They were asked whether they had chosen their child's school and if not, to say which type they would have chosen. The teachers were asked to give their sex,age, marital status and details of their education and teaching experience, and to state whether they thought that secondary modern schools should be mixed or single sexed. Both parents and teachers were asked to give reasons for their choice of schooling. All hypotheses relating to the choice of school were tested by the chi-square statistic.

IV.—CONCLUSIONS.

(1) Boys and girls at the mixed schools were strongly in favour of co-education. The boys at the boys' schools also preferred co-education, but the girls at the girls' schools preferred single-sexed schools.

(2) The older children of both sexes in all types of schools showed a stronger preference for co-education than the younger children.

(3) The majority of the parents who completed the questionnaires preferred their children to be educated in single-sexed schools.

(4) Men teachers with experience in both types of school preferred co-education.

(5) Women teachers with less than ten years' experience preferred mixed schools and those with ten years' or more experience preferred single-sexed schools.

(6) The children who preferred mixed schools considered the development of good boy-girl relationships to be their main advantage, while those who preferred single-sexed schools did so because they felt that the presence of the opposite sex had a disturbing influence and because they preferred to be taught by members of their own sex.

(7) The parents who preferred mixed schools did so because they considered that these schools fostered good relationships between boys and girls, provided good social opportunities and gave a broader outlook on life. The parents who preferred single-sexed schools did so because they believed the presence of the opposite sex to be a distraction from work and that there was poorer discipline in mixed schools.

(8) The teachers who preferred mixed schools believed that they develop better relationships between boys and girls, while the teachers who preferred single-sexed schools believed that there is a better standard of behaviour and discipline in these schools.

(*Manuscript received* 15*th February*, 1967)

AN INVESTIGATION OF PERSONALITY MEASURES ASSOCIATED WITH OVER AND UNDER ACHIEVEMENT IN ENGLISH AND ARITHMETIC

By L. W. RIDDING

(*Abstract of a thesis submitted for the degree of M.Ed., University of Manchester*, 1966)

I.—Aim.

The aim of the research was to use Cattell's and Eysenck's questionnaire methods of assessing common personality traits to investigate personality differences between pupils over- and under-achieving in English and Arithmetic.

II.—Method.

(*a*) The sample consisted of 600 boys and girls aged 12 plus, selected on the basis of test results obtained from the total entry for one year in Manchester schools. The sample was designed to provide for comparisons between over-achievers, under-achievers and average achievers. The first two categories consisted of pupils whose attainment scores deviated by at least half a standard deviation from their verbal reasoning scores. Separate groups were constituted for arithmetic and English and for both sexes. The total matrix consisted of twelve groups each of fifty children (two subjects × both sexes × three achievement groups).

(*b*) The children completed Forms A and B of Cattell's H.S.P.Q., yielding thirteen first order and two second order factors, and also the Children's Questionnaire adapted from Eysenck's M.P.I., which gives two measures—Neuroticism and Extraversion.

Analyses of variance on each of the seventeen personality variables and three Moray House attainment variables were carried out to investigate differences between (i) the twelve achievement groups and (ii) social class levels.

III.—Results.

(1) Eysenck's Extraversion was associated with a difference between the sexes. Girls exhibited more extraversion than boys.

(2) Significant differences occur between the mean score of Manchester boys and girls on eleven of Cattell's factors. Girls are more *anxious*, i.e., more tense (Q_4+) ; more unsure (O+) ; less stable (C –) ; less adventurous (H –) ; more *extraverted*, i.e., more sociable (A+) ; more surgent (F+) ; and less self-sufficient (Q_2 –) ; but more tender-minded and sensitive (I+) and less dominant (E –) than the boys.

(3) Cattell's primary factor ' G ' (Conscientiousness) differentiated significantly between English and Arithmetic, boys and girls in the Arithmetic groups showing more Conscientiousness than those in the English groups.

(4) The following factors showed associations significant at the 5 per cent. level:

Eysenck's Neuroticism: with over achievement in English. Girls show more Neuroticism than boys.
Eysenck's Extraversion: with under achievement in Arithmetic. Girls show more Extraversion than boys.
Cattell's primary factors E, F and G.
' E ' (Dominance): (i) Over achievers show more Dominance than average or under achievers. (ii) Boys show more Dominance than girls.
' F ' (Surgency): (i) Over and under achievers show more Surgency than average achievers. (ii) Boys: over achievers in Arithmetic are more Surgent than average or under achievers. (iii) Girls: over achievers in English are more Surgent than average achievers. (iv) Girls show more Surgency than boys.
' G ' (Conscientiousness): Girl over achievers in Arithmetic show more Conscientiousness than under achievers.

(5) Manchester pupils in this sample drawn from sixty-three secondary schools of all kinds excepting grammar show significantly less anxiety and introversion than those investigated by Butcher, Ainsworth and Nesbitt (1963), drawn from a single urban secondary modern school in Manchester.

(6) None of the personality measures showed any significant differences ocurring between social class levels obtained by grading occupations of parents on a modified version of the Central Register Office's Classification of Occupations, 1951 scale.

The main findings were that:

(i) No significant relationship was found between stability and over or under achievement.

(ii) Extraversion was found to be correlated with over-achievement.

(iii) No significant relationship was found between anxiety and over or under achievement.

(iv) Under achievement was found to be correlated with introversion. The child at 12 falls into the junior rather than the senior category of school pupils in respect of stability, although the analogous trait of conscientiousness (G+) was associated with over achievement in arithmetic.

It must be emphasised, however, (i) that this investigation is concerned with over and under achievement, i.e., with relative rather than absolute level of attainment (ii) that significant results were confined to the primary extraversion traits ' F ' (surgency) and ' E ' (dominance); (iii) that significant results were not obtained for every sex group or for every subject.

IV.—REFERENCE.

BUTCHER, H. J., AINSWORTH, M., and NESBITT, J. E. (1963). Personality factors and school achievement—a comparison of British and American children. *Brit. J. Educ. Psychol.*, 33, 276-285.

(*Manuscript received* 15*th February*, 1967)

THE IDEA OF GOD AMONG ROMAN CATHOLIC CHILDREN AGED SEVEN AND EIGHT

By MOTHER MIRIAM, S.H.C.J. (D. MARTYN)

(*Abstract of thesis submitted for the degree of M.Ed. of the University of Manchester,* 1966)

I.—AIM OF THE INVESTIGATION.

The aim of this research was to explore, using tests giving a maximum amount of freedom from suggestion, the content and kind of thinking which goes to make up the idea of God among Roman Catholic children aged 7 and 8.

II.—SUBJECTS.

A sample of fifty boys and fifty girls, with an age range from 7 years 6 months to 8 years 5 months, and an I Q. range from 70 to 136, was drawn from four Roman Catholic primary schools in the inner urban district of a city in the North of England. There were thus four variables—school, sex, I.Q. and C.A. With the exception of I.Q. ($P<\cdot025$) there was no significant difference in the four variables' contribution to the whole sample.

III.—METHOD.

A clinical interview was selected as the appropriate method. Four tests, two verbal and two using pictures, were constructed requiring the children to 'tell about God.' Throughout the tests projective technique in the form of dramatic impersonation of a missionary and role playing was used. A written verbatim record of each child's responses was made. The responses were divided into statements which varied in length from a phrase to a paragraph. The following categories and sub-headings were necessary for complete classification :

Category 1. Verbalism : (*a*) Old Testament ; (*b*) New Testament ; (*c*) Imperatives ; (*d*) Doctrine.

Category 2. Realism : (*a*) Physical Anthropomorphism ; (*b*) Action Anthropomorphism ; (*c*) Moral Anthropomorphism ; (*d*) Empathic Anthropomorphism ; (*e*) Personal Experience.

Category 3. (*a*) Magic ; (*b*) Fantasy.

Category 4. Religious Thinking : (*a*) Omnipotence ; (*b*) Omniscience ; (*c*) Omnipresence ; (*d*) Transcendence ; (*e*) Fatherhood ; (*f*) Prayer, worship and sacraments.

Each statement was allocated to one sub-heading within this scheme. The variables were examined using the chi-squared technique and correlation.

IV.—RESULTS.

The range in the number of statements offered by the children was 9—98 with a mean of 49·19 statements. A significantly larger number of statements ($P<\cdot001$) were classified as verbalism (31·12 per cent.) and realism (36·29 per cent.) than as magic and fantasy (10·44 per cent.) and religious thinking (22·15 per cent.).

Responses classified as verbalism and realism show that the children have no awareness of God as the Supreme Spirit. They think of Him as someone who is tangible, can be visualised, fits into their experience and has the properties of and behaves as a man. Their understanding of God is confused and partial, being limited by their level of thinking, which is essentially concrete. When the children cannot 'fit' God into these four ideas, they regress to pre-operational thinking expressed in magic and fantasy Crude literalisms, intuitive associations, partial and fragmentary thinking abound. The statements classified as religious thinking show a sense of wonder and awe, especially for God's omnipotence. There is a significant positive correlation between realism and religious thinking ($r=+0\cdot211$). As there is a negative correlation between personal experience and religious thinking (r= –0·271), the association must be between action anthropomorphism and religious thinking.

The following relationships with the four variables (with significance levels) were indicated :

(1) Level of intelligence is related to :
   (*a*) Number of statements given (·01).
   (*b*) Number of statements classified as religious thinking :
      whole category (·01).
      omniscience (·01).
      transcendence (·01).
   (*c*) Number of statements classified as verbalism (·05).
   (*d*) Number of statements classified as physical anthropomorphism (·05).
   (*e*) Number of statements classified as magic (·01).
(2) The school attended is related to :
   (*a*) Amount of verbalism given (·025).
   (*b*) Amount of responses classified as doctrine (·01).
   (*c*) Amount of statements classified as transcendence (·05).
(3) Chronological age is related to :
   (*a*) Number of statements given (·05).
   (*b*) Number of statements classified as realism (·05).
   (*c*) Number of statements classified as religious thinking (·05).
   (*d*) Number of statements classified as fantasy (·05).
(4) Sex is related to :
   (*a*) Number of statements classified as magic (·05).

Sections (1) and (3) tested by correlation : Sections (2) and (4) by chi-square.

(*Manuscript received 7th March,* 1967)

AN INVESTIGATION OF PERSONALITY ATTRIBUTES OF STAMMERING, ENURETIC AND SCHOOL-PHOBIC CHILDREN

By A. McHALE
(*The Leicestershire Schools Psychological Service, Leicester*)

(*Summary of a thesis presented for the degree of Ph.D. at the University of Leicester,* 1966)

I.—Aim.

The aim of the present research was to assess objectively the differences between groups of stammering, enuretic, school-phobic and ' normal ' children ; to discover possible evidence for the existence of a ' stammering personality,' an ' enuretic personality ' and a ' school-phobic personality.'

II.—Population used.

All the clinic cases had been referred to the Leicestershire Schools Psychological Service. The stammers had already been interviewed by the speech therapist with a view to treatment. The enuretics had suffered from enuresis for several years beyond the point at which dryness might be expected. The school-phobics had been absent from school for several weeks or had attended irregularly for a long period and showed the usual symptoms of vomiting, feelings of sickness and phobias associated with attendance at school. The diagnosis of each case was independently confirmed by the speech therapist, the psychiatrist, the psychiatric social worker or the school medical officer. Delinquents and children with low I.Qs. or from broken homes were excluded. The controls were selected at random from two junior and three senior schools.

III.—Tests.

Four tests were used in the research. The Terman-Merrill Intelligence Scale and The Vineland Social Maturity Scale were administered. Six needs were selected from Murray's list (Murray, 1938) as having particular relevance to stammering, enuresis and school-phobia, namely, aggression, dominance, autonomy, blamavoidance, infavoidance and harmavoidance. A measure of the 'admitted' needs was then obtained by means of a questionnaire derived from Murray's questionnaire material. The 'projected' needs were elicited by an open projective test consisting of three stories written by each child on any chosen topics. No pictures or stimulus material of any kind were used.

It was found that nine additional needs occurred in the stories with sufficient frequency to warrant investigation, namely, abasement, achievement, acquisition, affiliation, cognisance, exposition, nurturance, order and succorance.

The stories were scored in respect of four formal measures as follows : (*a*) the verb-adjective quotient (VAQ), i.e., the ratio of verbs to adjectives ; (*b*) the percentages of happy, unhappy and indecisive story-endings ; (*c*) the total number of expressions of all the fifteen needs ; and (*d*) the percentages of stories written in the past, present and future tenses.

IV.—Results.

Of a total of 170 children who were tested, 146 were retained in six groups as follows : stammering boys (St.) 24, enuretic boys (En.) 28, school-phobic boys (SPB) 22, school-phobic girls (SPG) 24, control boys (CB) 24 and control girls (CT) 24. The age range extended from 7½ years to 15 years, with a mean age of about 11¾ years. The mean age of the En. group was significantly lower than the mean ages of the other clinic groups. The range of I.Q. was from 82 to 171, with a mean of about 114. There was no significant difference between any of the groups.

Social maturity and the questionnaire results are given in Table 1, those for the projective test scores in Table 2.

TABLE 1

The Social Maturity Scale and the Questionnaire. Significant Differences Between Groups.

Variable	Groups	Significance
Social Quotient	CB exceeded SPB	++
Social Quotient	En exceeded SPB	++
Social Quotient	St exceeded SPB	+
Social Quotient	CB exceeded CG	+
Dominance	CB exceeded SPB	+
Dominance	St exceeded SPB	++
Dominance	En exceeded SPB	+
Harmavoidance	SPB exceeded St	++
Harmavoidance	SPB exceeded En	+
Harmavoidance	CG exceeded SPG	+

TABLE 2

The Projective Test. Significant Differences between Groups.

Variable	Groups	Significance
Infavoidance	CG exceeded CB	+
Harmavoidance	CB exceeded CG	+
Dominance	CB exceeded St	+
Harmavoidance	CB exceeded St	+
Infavoidance	CG exceeded SPG	+
Harmavoidance	CB exceeded En	+
Aggression	SPB exceeded SPG	+
Achievement	CB exceeded CG	++
Cognisance	CB exceeded CG	+
Nurturance	CG exceeded CB	+
Achievement	SPG exceeded CG	+
Abasement	St exceeded CB	++
Acquisition	St exceeded CB	++
Exposition	St exceeded CB	+
Cognisance	CB exceeded SPB	+
Acquisition	St exceeded SPB	+
Acquisition	En exceeded CB	++
Acquisition	SPG exceeded SPB	+

+Significant at the 5 per cent. level of confidence.
++Significant at the 1 per cent. level of confidence.

No significant differences between groups were found with respect to the verb adjective quotient, but on the story-endings, tense and mean number of needs, the SPB used the past tense to a significantly less extent than the CB. No significant differences between groups were found in respect of the kind of story-ending or the mean number of needs expressed.

The test results were drawn up into group profiles to highlight the main points of difference between the groups. It was found that each group had a characteristic pattern of results which was then interpreted in relation to the classifying symptom so as to present as coherent a picture as possible. The profile of the SPB group is given here as an example (Table 3). It appears to reflect the reported timid, withdrawn behaviour of school-phobics in situations outside the home and their failure to face up to external demands ; their lack of confidence and self-assertion ; and their weakness of social and intellectual drive and diminished school attainment. The less frequent use of the past tense may be regarded as a feature more appropriate to younger children and may be another indication that the school-phobic boys tended to be less confident and less mature than the controls. This profile seemed to be evidence that the school-phobic boys were characterized by a high level of anxiety, were probably the most maladjusted group and suffered from the highest degree of ' emotional disturbance.'

TABLE 3

Group Profile of the School Phobic Boys and Controls.

Direct measures : On the social maturity scale the SPB had lower (++) scores than the CB ; on admitted dominance the SPB had lower (+) scores than the CB.

Oblique projective measures :
Projected cognisance : The SPB had lower (+) scores than the CB.

Oblique formal measures : The SPB used the past tense less (+) than the CB.

The needs aggression, autonomy and blamavoidance failed to differentiate significantly between the clinic groups and the controls at either the questionnaire or the projective level. Explanations of these symptoms in terms of aggression received no support from the present findings. On the other hand, several needs which had not been hypothesized at the outset were found to differentiate significantly between the groups at the projective level, especially the needs abasement, achievement, acquisition and exposition.

V.—REFERENCES.

MURRAY, H. A. (1938). *Explorations in Personality.* Oxford University Press.

(*Manuscript received 9th February,* 1967)

OBJECTIVE INDICES IN THE ASSESSMENT OF ESSAYS

BY R. A. NEWBERY

(*Abstract of a thesis submitted for the degree of M.Ed., University of Manchester,* 1966)

I.—AIM.

The aim of this investigation was to consider the discriminating value of objective indices which might reveal a tendency towards more formal treatment of an essay, according to the nature of the topic and the increasing age (maturity) of the writer.

An attempt was made to consider :

(1) Whether objective indices which appear appropriate measures for the assessment of essays are reliable ; i.e., whether there is a high correlation between indices extracted from sets of essays on (*a*) a potentially ' concrete ' topic, and (*b*) a potentially ' abstract ' topic.

(2) Whether such indices reveal significant differences between the two sets of essays ; i.e., whether the essays on the ' abstract ' topic are measurably superior to those on the ' concrete ' topic, using objective indices as criteria.

(3) Whether significant differences can be found in such essays written by markedly different groups of students and, if so, whether such differences may reasonably be attributed to increasing age (maturity) or to topic.

II.—METHOD.

Part I.

In an attempt to explore problems (1) and (2) above, twenty 15–16-year-old ex-secondary modern school girls (the S.M. group) were each set two essays, one on ' The Cinema ' (' concrete ') and one on ' Juvenile Delinquency ' (' abstract '). Percentage incidences of subordinate adverbial, noun, and adjectival clauses were calculated, both separately for the two sets of essays, and for the total forty essays.

Part II.

The same two essay topics were set to :

(i) Twenty able 14–16-year-old grammar school girls (the G.S. group).

(ii) Twenty young women, aged 18 or over, at a college of education (the T.C. group).

(iii) Twenty post-graduate students (P.G. group).

It is reasonable to assume that all three of these groups were of superior intelligence to the SM group, and that the last were not only older but also more intelligent than any of the other three. For each of these groups, essays were examined with respect to the following indices, selected from the variables used for Part I of this study:

(1) Mean length of sentence.

(2) Mean length of longest three sentences.

(3) Percentage incidence of 'rare' words.

(4) Ratio of total number of adjectives, and substantives used as epithets, to total number of words used in each essay.

(5) Subordination index.

(6) Gross number of words used.

The same measures were obtained for the above six variables as were obtained for those studied in Part I

III.—RESULTS AND CONCLUSIONS.

Part I.

Of the correlations obtained for twenty-one of the variables examined, in the SM group's essays, nine were negative and twelve were positive. Only five produced values of r which exceeded $+\cdot 50$. It was suggested that these discouraging results might be due to the homogeneity of the sample.

Differences between topics were not significant for fifteen variables, but significant at the 10 per cent. level for one variable; at the 5 per cent. level for two variables; at the 2 per cent. level for one variable; and at the 1 per cent. level for two variables. 'Mean length of sentence' and 'Mean length of longest three sentences' were the only two variables which produced both a significant difference (at the 5 per cent. level) between the means for the two topics, *and* a correlation in excess of $+\cdot 50$.

A study of the distribution and type of subordinate clause usage suggested a number of variations between topics, but in many respects the frequencies of clause usage agreed with those found by previous investigators, presumably owing to the limitations imposed upon written language by its grammatical structure.

Part II.

Correlations between topics for the various groups, and for the combined samples, for the previously-listed six variables, produced values of r ranging from $+\cdot 71$ (variable (6), combined samples) to $-\cdot 21$ (variable (5); T.C. group), the most promisingly consistent positive values being for variable (1), whose r values ranged from $+\cdot 31$ to $+\cdot 68$,; and for variable (6), whose r values ranged from $+\cdot 27$ to $+\cdot 71$.

The following variables revealed significant differences at the 1 per cent. level when two-way analysis of variance was used: variables (1), (2), (3), and (6) between groups; variables (3) and (5) between topics; variable (4) interaction.

For variable (1), (length of sentence) between group differences showed the P.G. group to be superior to all others ($P < \cdot 01$), the differences between all other parts of groups being not significant. It was inferred that the P.G. group's superiority was a product of age and intelligence, since neither age alone, nor intelligence alone, sufficed to produce a significant difference.

For variable (2) (length of longest sentence), the P.G. group was significantly superior, followed by the G.S. group, the difference between the S.M. and T.C groups being not significant. It was concluded that variable (2) was not a satis-. factory substitute measure for variable (1), at least for the four groups studied.

Variable (3) (rare words), for all four groups, produced a somewhat higher mean for the 'concrete' topic than for the 'abstract' topic, possibly because of the artificiality of the criterion used (Thorndike's list). However, the differences were significant only for the G.S. and P.G. groups, and then only at the 10 per cent. level.

There was a more or less consistent progression in the percentage of ' rare ' words used, from the S.M. to the P.G. group, all differences between pairs of groups being significant at the 5 per cent. level or better, except for the difference between the S.M. group and the G S. group, which was not significant. These results suggested that if the criterion used is indicative of the extensiveness of a group's vocabulary, age (maturity) seems more important than intelligence in the development of ' rare ' word vocabularies. The criterion may well be valid when results for essays on two (or more) different topics are averaged, even if not valid for comparisons between essays on the same topic.

The interaction for variable (4) (use of adjectives) was attributed to the fact that, whereas the essays on ' The Cinema ' showed a steady progression towards the more frequent use of this variable, from the S.M. to the P.G. group, there was no such tendency in the essays on ' Juvenile Delinquency.' This difference may relate to the progressive qualitative improvement of the essays on ' The Cinema,' whose treatment tended to be factual and concrete at the S.M. group level, but progressively more general and abstract at G.S., T.C., and P.G. group levels No such progression was observed in the essays on ' Juvenile Delinquency,' although more mature ' insights ' were apparent in the essays of the T.C. and P.G. groups, possibly because of their psychological studies and their greater age and experience.

For variable (5) (subordination index), for all groups except the S.M. group, the ' abstract ' topic evoked a significantly greater use of subordination than the ' concrete ' topic. (For both the G.S. and the T.C. groups, $P < \cdot 01$; for the P.G. group, $P < \cdot 02$). The S.M. group's result (not significant) may have been due to the homogeneity of the sample. The mean index, for both sets of essays, was very similar for all four groups, suggesting that the overall use of subordination may have ' levelled off ' by about the age of 15–16 years, at least for normal writing purposes : other methods, namely the use of infinitives, gerunds, participles, and phrases, can be effectively used to express the same relationships.

For variable (6) (from number of words), the P.G. and G.S. groups used far more words than did the S.M. and T.C. groups ($P < \cdot 01$), the apparently huge difference in fluency being tentatively attributed to examination training.

Language ' quality ' is an elusive property, and does not always lend itself readily to a completely objective, quantitative analysis. It would seem that the apt choice, combination, and inter-relationship of words, to form a complex conceptual hierarchy, are the ways in which individual mastery of language is manifested. It appears, therefore, that while objective measures used in this study are all useful, something more is required for a precise evaluation of the quality of language usage, and of the formal or abstract nature of the concepts revealed in the writings of older students, than is yet available.

(*Manuscript received 5th December,* 1966)

CRITICAL NOTICES

The Schools Council (1966). *Examining at 16+. The Report of the Joint G.C.E./C.S.E. Committee of the Schools Council.* H.M.S.O., pp. 20, 2s. 6d.

By NORMAN FRANCE

(*University of Saskatchewan, Regina*)

The subject of this report is bound to generate strong emotional reaction at whatever level it is discussed. The Committee were well aware of the issues at stake and they appear to have steered a middle course clear of the inevitable controversies. They rightly emphasise that the distribution of attainment is continuous and attempt to reconcile this principle with the inevitable duality of the G.C.E. /C.S.E. examinations systems. The time, they think, is not yet ripe for the phasing out of the ordinary level of the G.C.E. Uniformity is to be sought instead through the description of the results. The method of achieving this is to be the subject of a programme of research and development. Much of the experience gained in developing the C.S.E. can be applied to the improvement of G.C.E. procedures. The internal mode of examining as practised in the C.S.E. and the greater use of subject advisory committees are but two examples of the kind of useful transfer recommended for the future. Even though . . . " Within the foreseeable future, there is . . . no question of abolishing the ordinary level G.C.E. examinations, or of altering their distinctive character," the encouragement by the Committee of the Mode 3 type of examining will pave the way for greater flexibility within a common pattern.

When the Committee is dealing with the grading scheme, two points of view are in evidence. On the one hand there is an attempt to preserve the concept of continuity but on the other hand, deep-seated traditional prejudices show through from time to time. For example, " It is a myth that the ordinary level pass standard is determined as a percentage of the entry . . . But even the most cursory examination of the figures for subjects with variable calibre entries . . . shows that variability of calibre is reflected in the results. Indeed, such variability has sometimes been a source of misdirected criticism of the work of the boards " (para. 66). This is a misleading assurance to say the least. In what other way is it possible to equate standards between two major subjects ? What interpretation could one legitimately draw from a difference in the failure rate between say, 45 per cent. and 35 per cent. ?

Some members of the Committee wished " to stop well short of creating a situation in which pupils received credit for doing bad work on inappropriate courses, assessed by unsuitable examinations. But they considered that there is, in fact, a narrow band of meritorious work, falling just below the threshold of an ordinary level pass, which could be identified and given recognition by means of a ' near-miss ' grade." (para. 58). . . . " We saw the introduction of a ' near-miss ' grade as tending to consolidate this unintended extension of the scope of ordinary level examining . . . Moreover, we did not consider that the introduction of a ' near-miss ' grade would solve the problems . . .it would merely create the new difficulty of drawing the line below which candidates just failed to achieve the ' near-miss ' grade " (para. 59).

The opportunity of achieving greater continuity by the extension of the grading system was thus given up in favour of the alternative of requesting schools to enter for the ordinary level of G.C E. examinations " only those candidates who are confidently expected to pass " (para. 60). The Committee rather naively believe that . . . " If this were done, the failure rate in the ordinary level examinations would fall dramatically. While it would be too much to expect that there would be no failures, we see no reason why a rate below 10 per cent. should not be achieved " (para. 60)

The theme of failure in grading and the reluctance to use that label recurs continually in any discussion of examinations. This is the root of all the emotional upset engendered amongst parents and teachers. What of the students themselves ?

Perhaps they are really less sensitive than we expect. An examination can suffer loss of prestige in the eyes of the candidates if there is no obvious differentiation between the ablest and the dullest. The time has come to face facts, to be more forthright in attitude and to be prepared to make an honest appraisal of what our pupils can do It is with this in mind that the following observations are offered

The basic purpose of any examination is to divide the student body into a number of groups. The groups will not be equal in size but the quality of achievement will rise in roughly equal steps from the lowest to the highest. In the simplest situation of all, there is only one group. In this case, the examination is open to all and all ' pass.' Such an examination has clearly little practical value At a slightly higher level of complexity, students may have to satisfy some criterion in order to be accepted as candidates. If then the results are expressed as a simple ' pass ' or ' fail,' there are essentially three groups. Unfortunately, in this case, the criterion for distinguishing between the two lower groups (fail/non-candidate) is different from that for the two upper groups (pass/fail). Herein lies the danger of the recommendation of the Committee that the schools should enter only those for the ordinary level G.C.E. examination who are confidently expected to pass. Teachers are not in a position to make this kind of judgment They can produce a reliable and valid estimate of the order in which their pupils should be arranged, but they cannot make a valid judgment as to where the line should be drawn in this order to separate the ' successes ' from the ' failures.'

We need to recognise that the labelling of a grade as ' pass ' or ' fail ' should not be the function of the examining bodies. If the ordinary level of the G.C E and the C.S.E. continue to operate side by side, then each will need to group their candidates in a series of numbered grades for which the overlap equivalents are very obvious to all concerned. The principle has been recognised in essence right from the inception of the C.S.E. and this Committee rightly emphasise that the two scales must be linked . . . " . . . both the schools and the candidates might be better served if all (G.C.E) boards supplied supplementary information on a common scale of not more than five grades, of which three would be passing grades . . . if the C.S E. boards wished to use the first three points on this scale to convey, privately, to the schools, similar information about the ranking of Grade 1 C.S.E. performances, we should not wish to raise objections " (para 55). If this suggestion were to be adopted, the scale of equivalents would be as follows

	G C E. ' pass '			G C.E. ' fail ' C.S.E. ' pass '			C S.E. ' fail '	
G.C.E.	1 (A)	1 (B)	1 (C)	2	3	—	—	—
C.S.E	1 (A)	1 (B)	1 (C)	2	3	4	5	6

A good deal of controversy would disappear if the G C.E. and C.S.E. boards confined themselves to allocating candidates to grades such as these. There would be full recognition of the fact of continuity, and this would transfer the responsibility for determining acceptable or unacceptable standards to those bodies charged with the duty of selecting suitable entrants to other courses of study or to employment This would also give much needed emphasis to the re-assessment of criteria for entry to certain kinds of employment. The Youth Employment Service would certainly welcome this, as too many employers present unrealistic demands unrelated to the qualities needed for later success. Close co-operation will be needed between the Central Youth Employment Executive and the research group of the Schools Council in order to produce the necessary detailed guide lines for employers in the interpretation of G C E. and C.S E. grades, and their use for predicting success in different occupations

The Report of the Joint G.C.E./C.S.E. Committee of the Schools Council should be widely read for the excellent review it provides of the problems involved in examining at 16+. Many psychologists will wish, perhaps, that the Committee had been rather more willing to advocate the changes in procedure now which will certainly be made in the long run. The need to give full recognition to the whole range of the continuous spectrum of ability in the school population is paramount. For the moment, the reactions of parents, of teachers and of employers are still to be reckoned with and the job of re-educating all concerned to a real understanding of the principles of examining is only just beginning.

DOWNING, JOHN (1967). *The i.t.a. Symposium: Research Report on the British Experiment with i.t.a.* National Foundation for Educational Research, pp. xvi+168, 25s.

BY GERTRUDE KEIR

In the past hundred years or so there have been a number of attempts to tidy up the spelling of the English language so as to facilitate reading and writing. It is safe to say that none have engendered more interest and controversy than the Initial Teaching Alphabet, devised by Sir James Pitman, and no other has instigated an educational experiment of such scope as the one reported upon in this publication. Although not final, for we are promised another this year, the report brings to a stage nearer completion the presentation of the results of the first British experiment and contains data relating to the use of the medium in the first three years of the follow-up stage. In it are given full details of the plan and procedures of the research, and the results presented supersede those already contained in the many interim reports of one kind and another which have appeared almost since the inception of the research.

The Symposium is divided into two sections. The first, by John Downing of the Reading Research Unit, the Institute of Education, University of London, is concerned with the research itself; the second records the evaluations of eleven experts, distinguished in the field of education. These evaluations are summarised at the end by Dr. Wall, Director of the National Foundation for Educational Research, and the whole is prefaced by an Introduction contributed by Lionel Elvin, Director of the Institute of Education—institutions both directly involved with the investigation from the start. As a separate addendum to the publication, there is a technical appendix which, we are told, may be obtained from the N.F.E.R. by those who wish to become acquainted with the data in more detail.*

As Morgan and Procter indicate in their evaluation, the reader who is familiar with the interim reports may have to adjust his mental set slightly when reading this one. The data presented here are somewhat different from those of earlier reports, for they relate to the results of matched groups of children, the statistical treatment is different, while the emphasis accorded to the various aims of the experiment has shifted somewhat, as might be expected in a longitudinal study. The references made by Downing to previous reports are minimal and, as has been observed, the presentation of the present results should have been integrated with those of previous publications.

Lionel Elvin points out in the Introduction, that the research first began in 1958, when the Institute of Education and the N.F.E.R. were approached by Sir James Pitman, who enquired whether an experiment could be carried out " to see whether children might learn to read more easily and effectively through an augmentation of our present alphabet instead of through our present very irregular orthography." Such an augmentation had already been devised by Sir James. Since his alphabet was designed, not to replace the present one, but for use only in the beginning stages of reading, the Augmented Roman Alphabet soon changed its name to the Initial Teaching Alphabet. Three characteristics are of importance · its augmentation by the introduction of special graphic symbols for sounds that may be a possible source of confusion for children, thus introducing regularity, the abolition of capital letters, and the design of the additional symbols in such a way as to facilitate visually the transfer to traditional orthography.

---

* Unfortunately, they may have some difficulty. The reviewer is much indebted to Mr. Latham of the Reading Research Unit for the loan of a copy for consultative purposes, it being impossible to obtain one from the N.F.E.R.

The aims of the investigation can be subsumed under two main headings, which can be put in the form of questions : I (1) Do children commencing reading learn more easily (more quickly, with greater accuracy, fluency and comprehension) by the use of i.t.a. than those learning through the medium of traditional orthography ? (2) What effects upon their progress (immediate and long-term) has the transfer to t.o.? II Since reading is closely linked with spelling and writing, are any differences to be found in the standards of their written work (as shown in length of composition, use of vocabulary and correctness of spelling) when compared with those of children taught by t.o.?

In order to answer these questions, matched groups of children were to be taught by the two different media, their progress being assessed at fixed intervals by means of basic reading books reached, and their results on standardised tests of reading attainment Tests of spelling and assessment of free compositions were to be carried out to answer the second point.

The original plan of procedure envisaged the participation of fifty infant schools with some 2,500 pupils taught through i.t.a to be matched with fifty schools with a similar number of pupils taught by t.o The schools were to be equated on those variables indicated by Joyce Morris in a previous N F.E.R. publication as being significant for the acquisition of reading skills in their pupils : location (urban or rural), type of school organisation, amenities of school buildings, number of pupils on roll, pupil/teacher ratio, minimum age of entry and general socio-economic level. When classes were later set up the last two variables were considered as referring to the actual age of the children and the occupation of their father.

The conditions were to be as rigorous and objective as the research team could ensure. However, one binding condition was that the scheme must be made as meaningful and valuable to the participants as possible, and must also obviously be conducted in a field setting. Field experiments usually mean the shifting of the ideal to the practicable and so it proved here. To start with, the number of Local Education Authorities willing to participate was rather small, particularly in the first year, as were the numbers of schools volunteering to use i.t.a. Matters improved in the second and third years during which intake of schools and pupils continued, so that by 1963 there were eighty-nine schools using i.t.a and sixty-nine using t.o., constituting a sample from which matching was possible. Of these, however, forty-eight i.t.a. and twenty-eight control schools could not be used. There was then left a pool of forty-one schools and 837 children in each group, from which the final samples could be drawn. These are large numbers but with the rigorous matching to be attempted, they are none too many Indeed, Downing himself points out that (certainly in the first year) matching on some of the school variables (school size, pupil/teacher ratio) could not be carried out with quite the rigour desired. School amenities also proved a little difficult to control. " The schools originally willing to participate in the experiment were too heterogeneous to permit a really good match to be made on *all* the variables listed." The research team were reasonably well satisfied with the matching, but it is worth noting that some 20 per cent. of the two groups were located in the same schools, and one wonders if this made any difference to the groups concerned. Further, in a few cases, the same school was included in the i.t.a group one year and the control group another year.

The matching of pupils was carried out with respect to sex, age, social class and intelligence, the latter assessed by means of Raven's Coloured Progressive Matrices and the Crichton Scale, administered as soon as possible after the child entered school by a teacher specially trained for that purpose. Matching of the main sample whose progress in the basic books was followed for the first year (N 651) shows that the experimental group was significantly younger at the 1 per cent. level of confidence than the control group. This age difference has disappeared at the next period of assessment (one term later), whether from natural causes or selection one cannot make out. It does show, however, that control of age by selecting September entrants is not as foolproof as it might seem, for 172 of the experimental as against 105 of the control group were below the age of $5\frac{1}{2}$ years (incidentally, the minimum age for which the Raven test is valid) Notwithstanding this age difference, the numbers of children scoring at or below the median (14) in the Raven's is closely similar for each group, suggesting that the experimental group might have been somewhat superior in intelligence. However, no answer can be given on this point, owing to the way in which the data are set out (in the technical appendix). These small points suggest that group matching on pupil variables was not so easy as was perhaps envisaged

What the experimental procedure did not set out to do was to allocate each matched pair of schools/classes randomly to the two groups. Nor was an attempt made to try and control the less tangible things which could be related to reading attainment (nor did the statistical treatment take these into account), one of the most important being the teacher's ability to shape learning. The second experiment has attempted to deal with this in its experimental procedure, unfortunately not with complete success. This comparative failure has, however, illuminated some interesting behavioural variables among teachers.* Method, too, was left to the individual teacher, as a matter of policy. Downing and his colleagues, no doubt rightly, considered that teachers should be free to teach through the new medium in the way they found most successful. This ties up with points raised by two of the symposiasts. Reid comments that the choice of a reading series, such as the Janet and John, may have penalised the control group, for as early as Book I nine different sounds for the letter ' o ' are introduced. The teacher, using a predominantly Look and Say approach in t.o. is likely to deal with these in one sort of way, the teacher emphasising phonics in t.o. in another, while the teacher using i.t.a. may rely chiefly on regularity of the system. However, it does seem that such a choice is justified, since it is one of the most popular of infant readers, and if the irregularity of English is a deterrent to progress, it is as well to know about it. A slightly different aspect of this question is raised by Gulliford, when he suggests that " for a complete evaluation of i.t.a. one would like to know from teachers how it has affected their teaching in practice—at what points procedures normal in teaching with t.o. have been omitted or reduced, introduced much earlier or in different ways " This would not only have thrown very interesting light on a feature of i t.a completely neglected in the Report, but would also highlight differences in method between teachers in the experimental and control groups participating in the research.

Another feature not controlled was the amount of time allotted to reading by teachers in both groups, as Vernon comments in her evaluation. Considerable attention, in contrast, was paid to the possible existence of the Hawthorne effect, and Downing describes the measures taken to deal with this, or to minimise it if it were operative. Downing was well aware of the added effects of the extensive publicity given to the research, probably inescapable from a situation where the interest and participation of teachers had to be actively maintained, and the co-operation of fund-granting bodies engaged, while at the same time conditions for the two groups were to be kept as similar as possible. Details of his efforts specifically directed towards the avoidance of the Hawthorne effect are presented in some detail, though their success is somewhat difficult to appraise. Reid, for instance, is of the opinion that the case for its non-existence has not been made out, only that efforts to simulate it were not successful.

With these general aspects of the procedure in mind, one may turn to the results of the teaching carried out through the two differing media. These may be considered at three stages in the experiment : the comparative attainments of the two groups (1) at the end of the first year, (2) just after the stage of transfer, and (3) during the third year.

At each of these stages standards of reading were assessed by means of progress through the basic readers, and results on the Schonell Graded Word Test and the Neale Analysis of Reading Ability. At a later stage the Standish Reading Test was also administered. To ensure adequate comparison, steps were taken within each pair of matched schools to equalise sample sizes, " so that the contribution of each member of the pair of schools to the total sample was the same." " For each critical test of attainment, subjects were discarded by random procedure to equalise the contributions of each school (or class) in the matched experimental and control pair." The comparability of the two samples was then checked in respect of any significant differences between the two groups in age, sex, social class and intelligence. Finally, the mean differences between the groups was estimated, and the statistical significance of the differences estimated by the Kolmogorov-Smirnov test.

---

* Downing, J. A , and Jones, B. (1966). Some problems of evaluating i t.a.—a second experiment. *Educ. Res.*, VIII, 100-114.

It is inevitable that there should be some wastage of sample during a longitudinal study. Downing himself states that " the numbers of children taking the different tests varied mainly because of the children's absences, but sometimes a whole class missed a test through some mischance in the testing programme in schools and this naturally has reduced the size of the sample. Sometimes tests were administered only to a sub-sample for reasons of economy." The sample was fairly intact at the time of the first assessment, end of first year (books and Schonell) and the fifth term (books, Schonell and Neale) ; hence these results may reasonably be accepted on that score. But only 291 children out of 484 (1961 and 1962 intake) contribute to the sample of test results on the Schonell in the mid-seventh term (mid-third year). The sample of third-year results available for the Neale tests shrinks in the same sort of way, being only 194 of the possible 322 (1961 and 1962 entrants) for whom results were reported at the end of the first year.* Hence, the results of the follow-up are much less clear cut. These samples were, of course, matched but, as Holmes points out, the selection of sub-samples " . . . leaves something of a cloud, one never knows how they would have performed had they stayed in the sample," while Reid comments that " it detracts greatly from the value of the results that so many different sub-samples have been taken from the initial ones. Under these circumstances, statistical inference is of very doubtful validity."

Downing has given the mean scores of the two groups, using the Kolmogorov-Smirnov test to estimate the significance of the differences between them. This has been questioned, notably by Holmes and Burt, the latter pointing out that this test has been specially designed to check *any* kind of difference in the two distributions. The consequence is that the method is apt to ascribe significance to a relatively small difference which would be rejected as non-significant by the customary procedure (calculating the standard error for each difference). Holmes stresses the point that " the sight of an estimate of an experimental error was completely lost. One cannot but wonder what the overall experimental error might be, when over a hundred separate significance tests were reported." Burt draws attention to the suitability of treating the data by analysis of variance. It is unfortunate that with the vast amount of data collected in the course of this long and costly study, the statistical treatment of the results should not have included a multivariate analysis. It is also unfortunate that rather little recognition seems to have been paid to the fact that statistical and practical (in this case educational) significance are not necessarily the same. Burt, indeed, stresses this in his evaluation, pointing out that with samples as large as these, quite small and unimportant differences can be shown to be significant in the statistical sense. But this fact is hard to accept when one is looking eagerly for meaningful results.

These criticisms should be borne in mind when the present results are inspected. So far as the first stage of learning is concerned, the progress of the i.t.a. group through the basic books was undoubtedly quicker than those children taught by t.o. This progress was maintained over two years of follow-up, the i.t.a. group being roughly a book ahead. The Schonell test results show a striking difference between the means of the two groups, particularly on the first two tests. At the same time, it must be pointed out that the mean scores of the t.o. group are very low, especially so at the end of the first year, a matter which merits closer examination. Indeed, the Schonell, like the Neale, seems to have been an unsuitable test, distributions for

* Data obtained from the technical appendix—figures relate to each group.

both groups being markedly skewed With the Neale test, the mean scores for accuracy are much higher for the i.t.a. group, though the differences between the i.t.a. and t.o. taught children for rate of reading and comprehension are small, being of the order of about 2 points in favour of i.t a (statistically significant). What one wants to know more about is the differential rate of progress made by various sorts of children over the period. Downing and his co-workers were extremely enthusiastic in some of the early reports about the use of i.t a for the slower reader. Graphs presented in this report, indicate that a substantial number of children in the i.t.a group made progress comparable to that of the t.o. group during the first year, and that an appreciable number failed to pull away from a low level of achievement during the next year. Hence, those teachers especially interested in the slow learning child may be somewhat disappointed by the evidence given in this report on its usefulness in this context. This is, of course, a result to be expected, when one takes into consideration the kinds of skills needed in learning to read and the individual differences amongst learners. No one kind of method or medium could really be expected to provide an answer to all the problems of the slow reader. There is evidence here, however, that some children do learn more quickly by the use of i.t a. and again one would like to know more about this.

That transfer holds up progress seems equally indisputable from the data presented here. Neglecting tables E6 to 8, where two-thirds of the i.t.a. children had not made the transfer, E9 clearly shows that children make lower scores after than before transfer, and this in spite of the fact that these 152 children, who had all been transferred at least six weeks prior to the giving of the test in t.o. must have been the brightest (mean scores on the i.t a. version of the Neale being higher than for the i.t.a. group as a whole). This suggests a period of readjustment considerably longer than anything envisaged in the light of previous statements on this problem, and lends weight to those who have queried what might happen to infants sent up to the junior school still using i.t.a. : or to the slower learner who might need very special help from a teacher not specifically trained to give it.

Having surmounted this hurdle, what then ? We turn to the results of tests in the middle of the third year (Schonell) and at the end (Neale). Only the 1961 and 1962 intakes had been followed through at the time when the data were being prepared for publication, a point already commented upon in connection with diminishing samples. The results presented here, moreover, refer to less than two-thirds of the 1961-62 intake, and they should be looked at with this in mind. The Schonell mean scores show that the gap between the two groups is closing, and again the difference between the means of the experimental and control groups for comprehension is of the same order (about 2 points difference in favour of i.t.a.) although the mean scores for rate and accuracy are about 7 points higher for i.t.a. taught children than for the t.o. group. These differential scores could suggest that teaching through i.t.a. emphasises certain aspects of reading, but does not substantially increase the child's power of comprehension, a result again borne out by the results of the one silent reading test used (the Standish) where indeed there is a slight difference in favour of t.o. taught children (two years taught).

Now these are interesting points, which awaken one's interest only to leave it relatively unsatisfied. A validation of the results reported here would be confirmatory evidence of differences found in similar, but better controlled, studies. Such a one is the second British experiment, to which Downing makes several references in the Symposium. The interim report upon this indicates that the schools affect successful learning to an extent not brought out in the first experiment. It also shows that when both groups are tested in t.o. there are very few schools in which the i.t.a. groups obtained higher mean scores. Indeed, of the four schools where differences reached a statistically significant level, three were in favour of the control group. Hence, one cannot but be somewhat doubtful about the general validity of the results reported in the Symposium.

So much for the results, which must be considered in relation to the question of sub-samples, of statistical test used and, as Holmes points out, of the actual number of tests of significance used in the light of the overall experimental error. The report gives an interesting account of some of the difficulties experienced in this sort of field experiment, but the tight frame of reference adhered to, however, has diminished the total value of the work. There are tantalising glimpses of problems, notably in the value of the i.t.a. medium for pupils of different capacities. The use of the mean scores obscures progress due to individual differences. What are the characteristics of children who can and cannot learn effectively through the new medium? As many symposiasts indicate, reading is made up of different kinds of skills, dependent on efficient functioning of, for instance, discriminative processes within the particular child. Perhaps i.t.a. may not be the most effective medium of change or as Downing points out in his tables on the errors made in the tests by the i.t.a. group, speed of transfer might be helped by the alteration of certain of the new graphic signs. Much more needs to be known about failure, rather than successes, for it is usually those who break down who give the research worker in learning most food for thought and provide the most varied and important clues as to what is taking place. It is to be hoped that this research, while not providing the answers to these questions, will act as an instigator for other, perhaps small-scale studies. Teachers participating in an active and informed way (as many have done, of course, in other studies) could get to grips with what really goes to make an efficient reader.

Although the symposiasts did not all agree in their comments, they were, however, unanimous in their praise of the energy, tenacity, enthusiasm and hard work shown by Downing and his colleagues in setting in motion and maintaining this very interesting experiment. In this, they will undoubtedly be joined by all readers of this book.

BOOK REVIEWS

DALE, D. M. C. (1967). *Deaf Children at Home and School.* London : University of London Press, pp. 272, 35s.

This book appears close on the heels of T. J. Watson's *Education of Hearing Handicapped Children* and, whether by accident or collusion, is an admirable complementary introduction to the problems of hearing impairment. Where Dr. Watson presents a well-structured reflective view of the oral method of education, Dr. Dale conveys more immediate experience, often at the personal and classroom level, resulting in a readable and interesting *pot-pourri* in the Ewing tradition, but in a lively, unpretentious, unpedantic and engaging style.

There is a wealth of practical wisdom, anecdote and example in this book which seems to be mainly directed at parents and teachers of deaf children. Parents will not easily find a more sympathetic and reassuring introduction to the particular problems of children with a hearing impairment and student teachers who may swear by others at examination times will surely come to bless Dr. Dale for this very useful guide through the difficult first years in the classroom.

In the main, this book is concerned with the highly specialised techniques of imparting language via residual hearing, lip-reading and speech. In common with Watson, an exclusively oralist position is maintained and no instruction by the manual communication methods currently used in many schools in the U.S.A. and U.S.S.R. and almost universally favoured by the adult deaf in Britain is recommended. Thus, the monopoly of oral methods in British teacher training long established by the Ewings in Manchester seems to be maintained in London. Oral methods unsupported by manual communication are not appropriate to the needs of profoundly deaf children without usable residual hearing and it is not surprising that this exposition of the current oral orthodoxy does not clearly define the word ' deaf ' which occurs in the book's title—almost always the word is used to mean partially deaf and those with usable residual hearing. Under the oral approach to education the needs of profoundly deaf children are largely ignored by the so-called ' genetic ' pretence that a deaf child has " the same needs and interests as any other child " an attitude which reassures parents and encourages teachers at the expense of profoundly deaf children, whose needs and abilities are very different from the loosely defined deaf group.

The section on signs and gestures quotes the basic tenet of the Ewings, which has yet to be validated, that signs inhibit the development of speech. We are reminded of the ancient custom of those grammar schools who prohibited English speech in the hope that boys would give tongue to classical languages uncontaminated by the unacceptable syntax of the vernacular. In fact, positive transfer based on much similarity of linguistic content ensures a positive correlation between English and Latin attainment. Similarly, positive relationships between manual communication and oral and written language skills have often been established in the literature, not excluding the bibliography appended to this work, but these findings are incompatible with an uncritical acceptance of the thoughts of the Ewings and no mention of this kind of deviationism occurs in the text: perhaps the book was written ten years ago.

A loose definition of deafness makes a nonsense of the few statistics presented in this book. We learn that gestures are tolerated in some oral schools for about 5 per cent. of the pupils who have great difficulty in lip-reading. Is this 5 per cent. of the partially deaf or the profoundly deaf, or the unfortunate few who are almost totally deaf ? Again, we are informed that two head teachers felt that 70 per cent. and 75 per cent. of their children could produce intelligible speech. These estimates are so wildly different from the sober 7 per cent. of profoundly deaf children with *fairly* fluent speech estimated recently by Scottish head teachers that one may excusably suspect that the two head teachers in question were either (*a*) assessing partially deaf pupils, (*b*) running small tutorial establishments, (*c*) insanely selective in their choice of pupils, or (*d*) self-deluded.

It is important that psychologists do not accept misleadingly optimistic estimates of the verbal communication abilities of deaf children as the misapplication of inappropriate psychometric instruments could, and does, lead to misdiagnosis often with quite tragic consequences. In general, this book is pragmatically accurate as an introduction to parents and teachers and it would be unfair to criticise Dr. Dale for a lack of cool objectivity in an area where positive optimism and confidence are more likely to generate effective results. But from an unimaginatively pedestrian point of view, this attitude can be misleading and after finishing this enjoyable and encouraging book, the reader might well be advised to have another look at Professor Tizard's foreword which sounds a jarring note of realism : " Many problems remain unsolved, and the majority of deaf children could still leave school better equipped for life in the ordinary community than they now do."

G. W. G. MONTGOMERY.

FAIRBANKS, GRANT (1966). *Experimental Phonetics*: *Selected Articles*. Urbana and London : University of Illinois Press, pp. 274, 49s.

Grant Fairbanks, a psychologist by training and for fourteen years Professor and Director of the Speech Research Laboratory at the University of Illinois, died in 1964 at the age of 53. An editor (who seems to be Dr. Murray Miran, a former student of Fairbanks) has collected in this book twenty-eight articles, five unpublished lectures, a complete bibliography, and even a list of graduate theses directed by Professor Fairbanks.

The unpublished articles are, generally, carefully reported accounts of carefully conducted research that might well be of interest to, and presumably already known to, researchers and specialists in relevant aspects of speech and hearing. Some of the topics investigated are : delayed auditory feedback: the effects of time compression of recorded speech on intelligibility: vowel intensities and vowel formants: a rhyme test for phonetic differentiation: vocal pitch in infants and children: articulation as affected by minor organic deviations of lips, tongue, teeth and hard palate.

A paper, originally published in 1954, that might still be of general interest is one in which a model is developed of the speech mechanism as a servosystem. By means of the now familiar notion of a feedback loop, what is going out of the system is compared with what is supposed to go out of the system so that what will go out of the system will be what should go out of the system. (What is supposed to go out of the system is not specified.)

With the exception of three and a half pages of previously unpublished " Additional analysis of the rhyme test," all articles are available in reputable journals for people pursuing lines of research similar to those of Professor Fairbanks. The editor has compiled a handsome, well-organised volume that few people will need.

COLIN FRASER.

GABRIEL, J. (1964). *Children Growing Up*. London: University of London Press, pp. 367, 30s.

This book is devoted to the study of children's personalities and written with the teacher in mind, for the author rightly believes that Education is concerned with the ' whole personality ' of the child in school. The material is arranged in four sections, part one dealing with the concept of personality, in which the discussion is based mainly on the theories of McDougall, Allport and Cattell. Having defined personality in terms of dispositional tendencies, hierarchical organisation and systems of needs and goals, Dr. Gabriel goes on to consider how these develop and the effects upon the maturing personality of environmental influences both cultural, parental and educational, as well as conditions related to temperamental and physical differences among children. In discussing the first group, he finds space to mention differences in child-rearing practices found amongst some primitive societies and different types of structures both in home and class room, e.g., authoritarian, *laissez-faire* and democratic, but omits reference to the differential effects of varying socio-economic and educational backgrounds. Though what he has to say about parental child-rearing practices is pertinent, his presentation is so condensed that some statements might mislead rather than enlighten. May it not be a little dangerous, for instance, to say so categorically that "in primary school, children of over-protective parents tend to be well behaved and to be prodigious readers, but they find difficulty in making friends and in participating in regular social activities with other children " (p. 79)? He does utter a later word of warning that the whole problem of parent-child relationships is a complex one, but most of what he writes about this is in a strongly positive tone. Similarly, in writing about the structure of the classroom, he has found himself forced to rely on broad general statements, and some of these may even worry the teacher. For instance, it is not only in a *laissez-faire* teaching climate that children write fantastic stories about voyages into space and so on. Now it is obvious to the experienced teacher that these sort of stories will also occur in the free writing books of the democratic school room, but the point is nowhere made explicit by the author. This is a niggling point, but serves to illustrate the rather hurried and compressed style of writing which dealing with a multitude of topics has impelled the author to use.

He does not find space, either, for any section dealing at length on the methods of teaching which are implied in the various kinds of structured teaching situations.

There are, too, a number of omissions in his chapter on the organisation of personality, although his presentation of this is clear, simple and interesting, covering briefly some learning theorists, as well as Freud, Sheldon, McDougall and Allport. But one would have liked some reference to the bearing of learning theory on the classroom situation rather more lengthy than is given.

Part three is devoted to children's emotional and social development. Gesell's work is considered, but rather brief references to that of Piaget. Dr. Gabriel's section on Play is interesting and useful, but here there is no specific mention of its importance in the Piagetian system. This whole section, however, is likely to be useful to the teacher. Perhaps the most striking section is the final one, in which he gives sound and simple rules by which the teacher can observe the child in school, and can chart both academic and personal maturity. His presentation here is full, interesting and helpful.

PENTROSE KENT.

GHISELLI, E. E. (1966). *The Validity of Occupational Aptitude Tests.* New York: John Wiley and Sons, pp. vii+155, 60s.

This is one of a series of books from the research programme of the Institute of Industrial Relations, University of California. It is a relatively small volume of some 150 pages, but its value bears no relation to its size. It is a mine of valuable information about the validity of aptitude tests for a wide spectrum of occupations.

The tests have been classified in five major groups : (1) Intelligence ; (2) Spatial and mechanical ; (3) Perceptual accuracy ; (4) Motor ; (5) Personality and interest.

Examples are given of the types of test material included in each category, but surprisingly, no examples of non-verbal intelligence tests are included in the first category. Hence, it is not clear whether the validities quoted for intelligence tests are intended to refer only to verbal intelligence tests

The book summarizes the whole field of American literature on the effectiveness of the different types of test in predicting success in jobs Each validity co-efficient quoted is followed by a code letter, indicating the size of sample on which the co-efficient is based. There are six categories corresponding to sample sizes ranging from less than 100 to more than 10,000.

The validities of the tests are discussed with reference to the prediction of two criteria of success (i) in training and (ii) in level of job proficiency actually achieved. On the whole tests are more effective in the prediction of trainability than they are of job proficiency (·30 compared with ·19). This is certainly true of tests of intellectual and of spatial and mechanical abilities, but tests of perceptual accuracy and of motor abilities predict trainability and job proficiency equally well. Dr. Ghiselli infers from this finding that business and industrial training programmes tend to stress the ' academic ' and ' intellectual ' just as much as education does in more formal school situations. As is to be expected, the discrepancy between the validity co-efficients for training and for job proficiency criteria is greatest in the case of trades and crafts and the industrial occupations

It is also interesting that the ability and trait requirements for job proficiency are quite specific to the job, whereas those required for training programmes have some measure of generality over a large proportion of jobs. This implies that most tests have some value in selecting for training courses whereas it is most important to use the correct test (or tests) for selecting personnel for an actual job.

There is an interesting chapter in which the organization of occupations is discussed in terms of ability requirements. It is concluded that there are three main dimensions of abilities in terms of occupational validity : one of intellectual abilities and perceptual accuracy, one of motor abilities and one of spatial and mechanical abilities, with the first two being positively related and the first being negatively related to the third. Thus, the validity coefficients of tests of intellectual and perceptual abilities correlated ·51 with those of motor abilities and —·10 with those of spatial and mechanical abilities

Dr. Ghiselli's book summarizes a vast amount of evidence from American studies that occupational behaviour is predictable by means of appropriate objective tests. He concludes that in most situations the tests have sufficiently high validity to be of considerable practical value in the selection of personnel.

I. MACFARLANE SMITH.

GOOCH, S., and KELLMER-PRINGLE, M. L. (1966). *Four Years On.* London : Longmans, pp.xii+167, 17s. 6d.

This book, the first of a series of Studies in Child Development from the National Bureau for Co-operation in Child Care, is a rather thin, disappointing, and in parts, badly written book. Altogether, one gets the impression that the authors were trying to make a book out of limited data obtained from too small a group of children. Though they do recognise that the limitations of their data justify only the most cautious and tentative of conclusions, they seem all the time to be striving to squeeze conclusions, albeit tentative conclusions, out of something all too thin. Their book is a follow-up study of eighty-one secondary modern school children aged 15 in 1964 (all of them pupils who had scored in the lower 60 per cent. of scores on the Neale Analysis of Reading four years earlier) selected from 262 boys and girls who had been extensively studied between the ages of 7 and 11 by means of a variety of intelligence, attainment and personality tests and questionnaires to parents and teachers. In the follow-up study three standardised tests, four projective measures, three questionnaires and teachers' confidential reports were used. Since two of the standardised tests had been used at the junior school level comparisons between earlier and later performances could be made. With three of the projective measures comparisons were made between what the children produced at 11 and 15. The only measures of general ability available were I Qs. from the Cornwell Intelligence Test administered at age 11.

Of the eighty-one pupils, thirty-seven had attended a ' traditional ' junior school and forty-four a ' progressive ' junior school. Since only fifteen of the boys from the ' traditional ' junior school went on to a ' traditional ' secondary modern school and the remaining boys and girls went to three secondary modern schools that were neither particularly ' traditional ' nor particularly ' progressive,' the authors recognise that this study cannot be regarded as a simple continuation of the original comparison of ' traditional ' and ' progressive ' approaches to education. Indeed, in their introduction they rightly state the general conclusion that questions such as " whether a child-centred or subject-centred approach achieves better results are so broad as to be almost meaningless." Nevertheless, they succumb at the end of the book to making the generalisation that " boys as a whole and the more intelligent girls do better in a traditional framework, while girls as a whole and the less intelligent boys perform best in a progressive framework."

One-eighth of the book is devoted to analysis of the results of the standardised tests. The authors realise the shortcomings of assessing scholastic progress by comparing scores on tests and re-tests, but the alternative model of progress which they suggest and use throughout the book is equally dubious: predicting attainment on the basis of measured I.Q. How justified is one in using a pupil's I.Q. computed at age 11 on a test designed for the age-range 8 to 11 to predict attainment at age 15?

The chapter on the projective material (one-third of the book) has some interesting material in it, but the reader is overwhelmed by the over-intensive analysis of limited data. Nevertheless, the study shows the value of longitudinal studies of projective material in assessing children's development, indicates that projective tasks may provide indices of stages in personal-social development and bear some relation to attainment level and rate of progress over the years, and suggests that they may prove useful in the study of children with adjustment and learning difficulties.

A. LAING.

GORDON, I. J. (1966). *Studying the Child in School.* New York: John Wiley, pp. xi+145, $3,85.

The author of this book is Professor of Education at the University of Florida, having previously worked at the Maryland Institute for Child Study. He thus has a combination of educational and psychological skills available to him, and it is this combination which he uses in this book with the admirable aim of helping teachers to understand and to apply in practice the latest developments in the field of child psychology. In particular, he is concerned to show that our recent realisation regarding the considerable modifiability of psychological growth processes and the dependence of intelligence on transactions with the environment have far-ranging implications for the teacher in terms of his role, his understanding of his pupils, and (above all) in terms of the teaching process. In addition to theoretical understanding, moreover, Gordon also proposes to hand the teacher a set of tools and techniques for the analysis of the learning situation and the improvement of teaching techniques, some of which refer to the intellectual and personality organisation of the learner, others to such aspects of the learning situation as the ecology of the classroom and the organisation of the peer group. Sociometry, creativity tests, Kagan's techniques for the assessment of cognitive style, Bruner's analysis of strategy, a self-concept scale, and (of course) Piaget's framework for intellectual growth—these are among the various techniques which are briefly explained and then bundled together in a kind of do-it-yourself kit.

To bridge the gap between the laboratory and real-life situations (in this case the classroom) is, of course, a most praiseworthy undertaking. Two objections must, however, be raised. In the first place, the terminology and level of explanation used is such that a considerable acquaintance with psychological techniques and theories is required if the reader is to benefit from the very brief expositions of the various approaches mentioned. A ten-page account of Piagetian theory, for instance, is unlikely to lead to any great insight, and the frequent introduction of unexplained technical terms like 'factor analysis' and 'phenotype' may simply give rise to bewilderment in many of the envisaged readers. The second difficulty stems from the author's choice of techniques, for amongst these are several that are still very new and hardly in a state where they can be recommended for diagnostic purposes. It may be unfortunate that the psychologist cannot provide more fool-proof techniques to the teachers, but it would be even more unfortunate if the teacher was led into precipitate use of methods on which a great deal of work needs still to be done before they can be applied outside a research setting. The book certainly performs the useful service of drawing the teacher's attention to the far more complex picture of intellectual functioning that is currently being adopted by psychologists, and also to the need to study the learning process in its social context (including the teacher's own behaviour). It should, however, be regarded as a progress report and not as a manual.

H. R. SCHAFFER.

GOSLING, G. W. H. (1966). *Marking English Compositions.* Victoria: Australian Council for Educational Research, pp. x+93, $2.50.

This reports on a system in which each child (15+) is given three topics of a different type: the synthesis of items of information (15 minutes); a description (15 minutes); a topic encouraging imagination and originality (45 minutes). No options are allowed. Each essay is marked by a different examiner who uses ten marked scripts as an aid to assessment on a fifteen-point scale. The markers also have the support of notes, briefing and practice sessions under the guidance of the Chief Examiner who is also available for second opinions. Each child's mark is a weighted sum.

The analysis of the results confirmed generally accepted conclusions: different abilities are assessed by different types of essay; there are ' halo ' effects in marking. Mr. Gosling also found that the variance from sexes of markers did not approach significance and that the time of marking affected the issue as variation between days rather than within marking periods. However, the chief interest of the report is the demonstration of the reliability of the system described. The interest of the administrator will no doubt focus on the practical aspects, including the need to run adequate pre-examination trials in order to provide the graded scale examples for the markers.

The book contains short and clear expositions of candidate and marker variability, halo effect, the effects of increasing the numbers of tests and markers and an examination of different methods of assessing reliability. There is a reminder that system reliability is not necesssarily guaranteed by marker reliability. This general survey, partly historical, is not as full as that in the Schools Council Examinations Bulletin No. 12—" Multiple Marking of English Compositions," which, for instance, clarifies the distinction between ' rater ' and ' colleague ' reliability.

These two publications form useful complements to each other.

D. YOUNG.

HILLMAN, J. (1967). *Insearch: Psychology and Religion.* London: Hodder and Stoughton, pp. 117, 21s.

The American-born author of this volume has a doctorate qualification in psychology, is a trained Jungian analyst and is currently Director of the C. J. Jung Institute in Zurich. The book is a four-part essay on varied subjects involving the relationship between psychology and religion and appears to be addressed to priests and ministers in their role as counsellors. It may be that the author's experiences in Switzerland and the United States have made him over-estimate the role of clergymen, but in Britain where there is great interest in religion, with a curious hostility to the organised church, the clergyman's role as counsellor is more doubtful.

The language of *Insearch*, as the title suggests, is somewhat mystical, imprecise, filled with Jungian archetypal metaphors and difficult to read. One statement leads to another with little evidence in support, although pearls of wisdom abound. Thus, " human problems are not something people have, but something people are," is understandable, as is the warning to the counsellor that curiosity can be a dangerous dissecting and impersonal process. On these grounds Dr. Hillman makes a side blow at psychological testing and in his argument for allowing a patient ' psychic space ' appears to be pleading for the Rogerian non-directive therapy as the ideal.

The religious element of this book is to be seen most clearly in chapters two and three where the ' soul ' is discussed as the reality behind the human's unconscious experience, and ' the problem of evil ' is seen as the inner darkness of the unconscious Counsellors are warned that it is the cure of souls that matters, not the second best substitute of the mind. All this is couched in existential prophetic-type writing, which at its best is poetic, but at its worst is totally obscure. Thus, we have to " befriend the dream " rather than dissect or understand it, and " who lives in me ? what inscapes are mine ? "

After attacking formal Freudian analysis, the author turns to assault the 'new theology' because a de-mythologised God is "a God disembowelled of emotion, a mental figment without psychic reality." In the final chapter the religious spirit is seen as the female anima as part of the human's bisexuality, and several delightful female archetypes are examined.

If the discipline known as the Psychology of Religion is to gain acceptance, this is certainly not a book to help such a process. It typifies the anecdotalism of armchair psychology of the last century, and provides no quantitative objective judgments which the work of the last ten years in the sphere of religious concepts, attitudes and beliefs has made possible. Significantly, the book has no bibliography.

RONALD GOLDMAN.

KELLMER-PRINGLE, M. L. (1967). *Studies in Child Development. Adoption Facts and Fallacies.* London: Longmans, pp. x+251, 21s.

The passage of the Children and Young Persons Act, 1963, authorised local authorities and the Home Office to conduct or support research in child care. Wishing to draw together and assess the growing volume of information on this subject, the Home Office approached the National Bureau for Co-operation in Child Care for assistance. The present publication is the first of a series giving the results of their investigations. Concentrating on various aspects of adoption, it seeks to take stock of what is known. Thus, three-quarters of the text is in the form of a compilation, summarising and listing numerous researches and studies. These are mainly drawn from work in U.S.A., Canada and the U.K., between 1948 and 1965. Three short introductory chapters by Dr. Pringle state the aims set and the methods used, and discuss the facts and fallacies unearthed. As to the first, she finds three reasonably attested: (1) The desirability of early placement, but "there is no agreement as yet as to what is an early or optimal age"; (2) the importance of the personal qualities of the adopter and his attitudes and feelings towards adoption; "the age, income and social class of the adopter are far less important"; (3) the difficulty that many adopters have in telling the child of his status but "there is need to investigate the most appropriate time and method of disclosure." Far more findings are considered 'not proved' or fallacious and this is not surprising considering the sketchiness of many of the researches.

In the chapters (4–8) reporting on the researches, a sort of hierarchy appears. Chapter 4 is devoted to considerable abstracts of completed research projects. arranged in date order under six heads (predictive studies, comparative studies, clinical studies, studies connected with administration and practice, community attitudes and unmarried mothers). Full summaries covering the aims of the researches, the nature of the samples, the methodology and the results, are given. Chapter 5 is an annotated bibliography of lesser studies, arranged chronologically, under four heads Chapters 6 and 7 give annotated lists in alphabetical order of on-going projects in U.S.A., Canada and the U.K. Finally, following three short appendices, is a composite bibliography, bringing together all the material reviewed in chapters 4-8

A book such as this stands or falls by its usefulness as a work of reference. Clearly, the present volume is a unique source of information on an important topic The abstracts and summaries prepared by an able team under Dr. Pringle's direction are well planned and of high quality. Their grouping is described above and this gives some assistance in locating different aspects of the general subject but it provides far from sufficient guidance. In fact, the absence of a subject index is the book's only weakness, but it is a serious one.

O. C. SAMPSON.

KELLMER-PRINGLE, M. L. (1967). *Studies in Child Development: 11,000 Seven-year-olds.* London: Longmans, pp. xx+246, 21s.

This is the first report of the National Child Development Study (1958, Cohort) and is an expansion of the evidence prepared for the Central Advisory Council for Education. In fact, apart from some medical evidence, most of this information has already appeared in the Plowden Report, Vol. 2, pp 401-543. It is the work of a team guided by a Steering Committee and a Consultative Committee, containing a number of distinguished persons In the foreword by the Chairman of the Steering Committee, it is stated that this report is " inevitably very much of an interim nature." It is further admitted that there are " marks of haste which will be apparent to the attentive reader." This report is yet another of these large scale surveys of every aspect of development which could conceivably be of importance Impressive though such data are at first glance, one finds on closer analysis that the thoroughness is in the number of items assembled on an extensive number of children, and that on any specific topic the relevant information is disappointing. Accepting the statement that this is an interim report, one is still left wondering why it was felt justified (even if necessary) to publish at this stage. To give this as a reason for quoting only the barest of facts does not seem adequate; while the frequently repeated excuse of 'lack of time' and 'largely a shortage of time' is tiresome to the reader and unforgivable in a published document.

The children who form the basis for the report are that proportion of an earlier study on maternity services and perinatal deaths which could be traced (original sample 17,000). Children from Scotland and Wales are excluded. Not only are there nearly 2,000 untraced, but there are others whose forms were not returned by the deadline set, and who are, therefore, excluded. The figures in the various tables vary markedly depending on the aspect being considered (7,986 on speech, 10,596 on reading schemes and 6,876 in the test and 10,808 in the appendix on reading ability as assessed by the teachers) Further, those more limited samples are not necessarily representative of the total group In one instance, as a result of " technical difficulties which could not be resolved in the time available " 300 punched cards could not be used in the analysis In short, it does not require an 'attentive reader' to observe the deficiencies of the report.

On the positive side, however, there is a great deal of factual information on the incidence of various abnormalities in a 7-year-old population—particularly on the medical side. On psychological aspects, the results tend to confirm previous work, though they do in some instances bring the figures of incidence up-to-date. There is some interesting information on sex differences. As in other studies, the girls were superior in reading, but the boys were found to be superior in arithmetic; the girls were classed as superior in creativity, according to the teachers In nearly all instances, there was a greater incidence of disturbance in boys than in girls—with the interesting exception of travel sickness which was commoner in girls.

Criticism of this presentation should not be taken as condemnation of the value of the long-term project Provided that the study goes beyond the amassing of more and more data, to intensive study of significant aspects, then one will await with interest a definitive statement based on a large and representative sample studied by an interdisciplinary team MARGARET M CLARK.

KLAUSMEIR, H. J., and HARRIS, C. W. (Eds.) (1966). *Analyses of Concept Learning.* New York. London: Academic Press, pp. xiii+272, $8.50.

BILODEAU, E A (Ed.) (1966). *Acquisition of Skill* New York. London: Academic Press, pp. xiii+539, $12.50.

There was a time, not long distant, when it was difficult to find a text in psychology, of American origin, which was not geared to the large undergraduate market. Now, thanks to grants for research, available from such diverse Federal agencies as the Office of Naval Research, the Surgeon General's Office, or the Office of Education, it is possible to organise research symposia for the interchange, on a personal basis, of the ideas of many workers in a given research area. The result is bound to be stimulating. With relatively short delays in printing, the results are available for a great many others who are likely to have research interests in these areas. By English standards the prices are high, but then so is the quality of the thought available.

A recent review in this *Journal* (XXXV, Pt. 1) noted of the Michigan symposium on the Psychology of Human Learning, that, " amid so many over-simplified accounts of human learning, one can welcome a book of this high-level kind." If one adds the present two symposia, as represented by the two books here reviewed, the Tulane Symposium on the Acquisition of Skill, and the Wisconsin Symposium on the Analyses of Concept Learning, the reader will be possessed of a worthy trilogy. Only one psychologist, Underwood, participated in all three symposia—and but two others participated in both the Michigan and Wisconsin meetings

The Wisconsin symposium was sponsored by the Research and Development Centre for Learning and Re-education, a centre established " to secure more efficient learning for children, youth, and adults, particularly related to concept learning, problem solving and the nurturance of related cognitive abilities," and this symposium was the first public venture of the centre to bring knowledge to bear upon concept learning. Four main topics were discussed: schemes for classifying and relating concepts, the learning of concepts, learning-teaching processes, and concepts in various subject fields. There were two participants from England, Harré, in the first, and Lovell in the last of these.

Merrifield does an excellent job of relating concepts to Guilford's model of the structure of intellect, and incidentally helps to fill in some of the gaps in the transition from Guilford's earlier two dimensional model, to his present three dimensional model. Seven psychologists are represented in the discussion on the learning of concepts. Underwood, as usual, is eminently readable and fundamental, Gagné makes a clearer delineation of concepts and principles than he does in his own text, but personal preference is for the paper by Kagan on the developmental approach to conceptual growth, closely followed by Jensen's paper on individual differences in concept learning. Between them, they provide a wealth of ideas for research.

There is no doubt that, in the section devoted to learning-teaching processes, Ausubel provides the outstanding paper. He is much less attached to the idea that all learning should be ' discovery ' learning, one of the residues of the Progressive Education Movement to which lip service is continually paid, and has a belief in the efficacy also of reception learning. Meaningful learning is perceived as the hub of the educative process, and this is related to subject matter mastery. He prefers to speak of concept assimilation, and of the cognitive structures to which new meanings can be anchored. A stable cognitive structure allows accurate and unambiguous meanings to emerge, an unstable one inhibits meaningful learning and retention Unfortunately, only a small part of the paper deals with ways of controlling the cognitive structure, being concerned mainly with the role of discriminability in meaningful reception learning.

The papers in the final section, whilst interesting, are less relevant to the underlying problems, but are indicative of attempts to apply current knowledge to problems associated with subject matter learning.

Whilst Klausmeir and Harris, by limiting their symposium to one major, and three minor sections, left themselves with a relatively straightforward editorial task, Bilodeau set himself a much more difficult task. The 539 pages make it the first full length book coverage of the topic since the acquisition of skill was lifted from the Bryan and Harter studies of the horse and buggy days to a position of importance as recognised by its inclusion as a separate topic, both in the Annual Review of Psychology and in the Michigan symposium.

Irion has given the book an excellent start with a brief (37 page) history of the topic and a ten page bibliography. Selective learning and individual differences are adequately covered in the next two chapters, then follow four chapters largely devoted to aspects of motor learning, and a further, exclusively motor, on tracking behaviour. The volume concludes with a paper by Underwood on relations and parallels between motor skills learning and verbal learning.

In terms of educational research two papers, in addition to Irion's historical account, are of particular value. Smith who has generally had a radical view on learning matters, continues with an excellent analysis, points from which might be most relevant in producing new theories about the process of teaching. Underwood argues that Smith's approach is too radical for this stage of our knowledge. Nevertheless it is still a chapter, partly speculative, worth several re-readings. The other important paper, from an educational viewpoint, is that of I. M. Bilodeau. Her work gives a most lively, succinct account of the literature on information feedback. If research skill is in any way related to knowledge of the literature, here would be an excellent point of departure.

The other chapters are less relevant to educational psychology, but are distinctly representative of the levels of skill, knowledge and sophistication necessary in psychology today. The book is a tribute to Bilodeau's avowed purpose—but it is not a book for the average educational psychologist, or the average educator.

REGINALD EDWARDS.

LYONS, J., and WALES, R. J. (1966). *Psycholinguistic Papers*. Edinburgh: University Press, pp. 234, 42s.

The interdisciplinary science of psycholinguistics is new. It may be said to have arrived on the scene in 1954 when Osgood and Sebeok published their monograph. The present volume offers a picture, both of the progress that has been made, and of the problems preoccupying the science today. It is a glimpse of work in progress, and its strengths and weaknesses stem from this fact.

The book contains the proceedings of the Psycholinguistics Conference held in Edinburgh in March, 1966. There are five main papers, by J. P. Thorne, E. T. Wales and J. C. Marshall, David McNeill, J. Foder and M. Garrett, and E. S. Klima and U. Bellugi. Each is followed by two prepared critiques by other students of the field, and by a summary of general discussion. The critical comment is mostly sharp and thorough, so that the overall impression is of unsolved problems rather than of results achieved. Since this is a record of a debate between experts, no concessions are made to the naive reader, who will find himself plunged into complex disputes carried on in a strange jargon. If he is prepared to work at it, and follow up references, the book could serve as a useful introduction to the subject ; but without such determ nation, the general reader would be advised to go elsewhere for his first introduction to the subject. For psycholinguists themselves, it will, of course be essential reading.

Possibly the major theoretical step taken during the last decade is the distinction made by Chomsky between competence and performance. Fodor and Garrett devote their paper to a discussion of it, and of the empirical evidence relevant to it. The central problem underlying their paper, which is brought out more clearly by Professor Sutherland's lucid comments on it, is just how a distinction drawn from linguistics can be paralleled psychologically. Briefly, by competence is usually meant the capacity of an idealized speaker-listener to generate all possible grammatically correct utterances within a language ; by performance is meant the utterances of actual people. The linguist's contribution is to analyse the grammatical structure of language so as to bring out those rules of phonetics, syntax and semantics which are implied in such a competence. The psychologist investigates linguistic performance and, as Sutherland says, attempts to infer mechanisms to account for this performance. What is problematic is how the grammar of language is relevant to psychological theory construction.

Perhaps the most impressive paper in the book is that by Wales and Marshall on linguistic performance. They give a masterly review of the theoretical problems and the existing empirical findings. Unfortunately, it is too rich and condensed for easy reading. The main purpose of the paper is to recommend a particular conceptual schema for analysing the stages intervening between linguistic input and output. Central to their schema is the distinction between surface and deep analysis. This corresponds to the linguist's distinction between surface grammar (the actual phrasing of sentences) and deep or basic grammar (the structure of the meaningful relationships between elements). The same deep structure can be expressed in varying surface structures, and different deep structures can have the same surface structure. As Wales and Marshall point out, this feature of the grammar of language—and therefore of the functioning of language users—poses considerable difficulties for orthodox learning theories, and even for the more sophisticated mediational extensions of these theories.

For students of child development, one of the most interesting papers is that by McNeill. Starting with the distinction between surface and deep structure, he sought to discover whether the latter was discernible in the earliest grammatical utterances of children. Reviewing data on American, Japanese and Russian children, he claims that certain basic grammatical relationships (such as ' subject-verb ') are present from the start, and under conditions which make it seem highly unlikely that they are a function of imitation and social learning. He ends with the speculation that the child's capacity to utter sentences consistent with deep or basic grammar is an expression in language of a more general cognitive ability, and one which is part of the innate equipment of language-using organisms. Many obvious, and some less obvious, comments are made on this paper by Colin Fraser and Margaret Donaldson.

The remaining papers include a report on the results of a study of the grammar of young children's speech, and a semi-philosophical paper on the ambiguity involved in ' hearing sentences ' by Thorne.

This is an exciting, if sometimes frustrating, book. Running through it are several focal problems of which three seem to stand out. The first is the relationship between formal grammar and the theoretical models of psychology (reminiscent of the problem Piaget faced in attempting a psychological analysis of logical thinking). The second is the problem of specifying psychologically the functions implied in the existence of deep structure in sentences. And the third is the specification of a psychological model to account for the transformation of deep into surface structure. DEREK WRIGHT.

MACLEAN, I. G. (1966). *Child Guidance and the School.* Methuen : Education Paperbacks, pp. 117, 9s. 6d.

This book has been written " to increase understanding of the problems of the normal and abnormal child, particularly in the forms in which student teachers or psychologists, teachers or parents may encounter them." It is clear that the author has mainly teachers in mind. Indeed, the title could very fittingly be changed to ' Educational Guidance,' for she makes the point that, with the shortage of psychologists and the few numbers of teachers specially trained to deal with the handicapped and maladjusted child, it is likely to be the ordinary teacher who has to deal with the psychological problems of learning and school adaptation of these children.

Consequently, it is of the utmost importance that teachers should be able to help their pupils by recognising when children need special help, by knowing where to seek it, if available, and by knowing how to cope with difficulties within the framework of school and classroom. Understanding the nature of normal and of deviant development, and the art of classroom management, can lead to the amelioration and prevention of difficulties. The knowledge, too, of what skilled assistance means is of paramount importance for teachers to acquire. Dr. MacLean's task is to introduce these matters to their attention and study.

She begins with a description of the Child Guidance Service, and this is likely to prove useful to all classes of readers, teachers, probation officers and others who work with children, as well as parents. The content covers the functions of the child guidance team (it might be pointed out that the author appears unaware that in England, as in Scotland, teaching is a prerequisite for training as an educational psychologist). The numbers and types of children likely to be in need of referral are also discussed (there is an obvious misprint on p. 16 where she states that from 1,000–1,500 would normally need to be referred from a school population of about 56,000, representing 0·03 per cent.) and she describes the procedures for examination and treatment.

She begins her consideration of handicapped groups by treating normal development in home and school, and the role of the teacher in fostering it. Enforced-compression, here as elsewhere, has led however, to some statements liable to be misconstrued. For instance, when referring to the distribution of intelligence (p. 18) " the normal or average are grouped towards the centre of the distribution and are usually the greatest part of it." Or when, in giving a brief history of educational provisions for the mentally handicapped, she states that " those (children) who were mentally defective were distinguished from those who were not by Act of Parliament " (p. 35). In chapter 3, the reader is made familiar with the needs and problems of the child with special handicaps, intellectual, educational, those of general physical handicap and those connected with specific disabilities—the partially sighted and hard of hearing. She deals with a great variety of conditions and their concomitant problems, together with the educational provisions which have been set up to deal with these, for instance the Occupation (Training) Centres, Special Classes, Opportunity Classes, and the like. The material has had to be greatly condensed and in consequence the treatment tends to be somewhat bitty and various statements are either misleading or open to further discussion. For instance, from the half page devoted to epilepsy, the depressing impression might be gained that the epileptic is nearly always mentally defective or shows severe behaviour disturbance, although the author would certainly be the last one to wish to convey such a notion. Again, not everyone might agree with the statement (p. 44) that " it is important for the teacher to remember that no matter how poor educational progress may be, a child who is ' dull ' in intelligence is not suitable for transfer to a school for mentally handicapped children." Much might depend, for instance, on the other characteristics of the child, or the home background, but space obviously does not permit a fuller discussion of this. The same could be said of some other topics treated in this chapter which, as a whole, is of interest. For instance, when describing various types of children with visual defects, and how the teacher can be helped to perceive these, she comments that " inability to distinguish colours may also give the teacher a clue " (p. 60). This may well leave people other than the teacher a little puzzled. Again, it is questionable whether the space devoted to laterality, rather more than half, in her four page discussion of specific reading disability, is likely to be of great help to teachers dealing with the majority of slow readers.

Dr. MacLean has selected some interesting topics to mention, the adopted child is one, and her chapters on the influence of the home on behaviour problems and the place of the teacher in treatment and prevention alike will certainly be of value. She has attempted a difficult task, which has meant the selection and compression of many important topics into a small volume. She has not always been successful in overcoming difficulties created by this situation. Nevertheless the book is likely to supply a want and many teachers could well use it as an introduction to a fuller study of children in need of guidance, whether in home or school.

GERTRUDE KEIR.

MORRIS, JOYCE M. (1966). *Standards and Progress in Reading.* London: National Foundation for Educational Research in England and Wales Research Reports. Second Series No. 1, pp. xxxiv+493, 50s.

As this publication contains the findings of various interrelated investigations that proceeded in Kent over some ten years, it merits close attention. Much of the material, however, is an analysis of test results obtained between 1955 and 1957 with the 1946 age group, and may be out of date. A Review of Recent Research in an Appendix by B. S. Cane, indicates that standards are probably improving generally. Nevertheless, teachers and administrators would find much to interest them if they were to make the effort to read and consider the implications of the somewhat prolix test.

The most interesting results relate to the study of children in ten selected schools classified as Good, Bad, Improver or Deteriorator on the basis of an earlier (1954) investigation. These last two such classifications were developed from cross sectional trends rather than longitudinal, which would have been better. A convincing case, however, is made for the representativeness of these schools in relation to schools in Kent. The characteristics and environments of Good and Poor readers in these schools were studied intensively between 1955 and 1957, ' Good ' and ' Poor ' being determined by standardised test results and stage of reading scheme reached. Those classified as Poor, about 14 per cent. as compared with 19 per cent. in 1954, were found likely to stay so, and it is estimated that about 7 per cent. of the Kent sample left school in 1961 still experiencing considerable difficulty which would probably handicap them in adult life. Such children tended to have more emotional difficulties, in terms of the Bristol Social Adjustment Guides, lower non-verbal I.Qs. and poorer socio-economic backgrounds than the Good readers. All this is hardly a revelation. What emerges of real importance is the personality and training of the teachers, those teaching Good and Poor readers being differentiated by interest in reading research, pupil-teacher relations and length and quality of training. In a situation where half of the 7-year-olds still required infant reading techniques, 75 per cent. of their teachers had not had appropriate training. A sound case is made for ensuring that teachers in primary schools and lower streams of secondary schools should know about the teaching of the basic mechanics of reading and, with the continuing shortage of educational psychologists, the diagnosis of difficulties.

Further ' Extensive ' studies involved 1,848 children between 1955 and 1957, and 2,253 children in 1957. Somewhat depressingly, a positive relationship was found between good reading and a formal approach to the teaching of beginners. Girls read better than boys at age 8, but the difference was not significant at the upper-Junior stage. The phonic *versus* whole word debate is, in this case, slightly resolved in favour of the latter. Neither method produced a higher level with beginners, but those taught by the whole word method showed " a greater rate of improvement."

Psychologists may be disappointed that no really new information emerges for example on the perceptual characteristics of the Good and Poor readers. The whole research is, however, a distinct contribution to the literature, although not the major one it might have been. G. H. TURNBULL.

ROSE, GORDON (1967). *Schools for Young Offenders.* International Library of Criminology: Delinquency and Social Behaviour No. 17. London: Tavistock Publications, pp. ix+244, 52s.

The history of the Approved School system in this country, and of the reformatory and industrial schools which preceded schools " approved by the Secretary of State," is an example of the strengths and weaknesses of our eclectic approach to educational and penal institutions. Less than a quarter of approved schools are run by local authorities, the rest by voluntary agencies, yet the law has intervened peremptorily in the lives of those who attend them. They receive not only children who have been guilty of an offence, but also children who have suffered one. A major function of these schools is to provide an education for children, many of whom are still of compulsory school age, but it is not provided within the educational system.

This book makes a timely appearance, and none the less for the delay in its publication which took place when it was expected that the Statutory Rules governing approved schools were to be revised. Shortly before its appearance, the Home Office White Paper, " The Child, the Family and the Young Offender," was published and Mr. Rose revised his manuscript to take account of its proposals. Now that the whole matter is in the melting pot again, his book is most welcome, and its readers will be grateful for an able summary by one who knows the field extremely well.

In the first half of the book the author has compressed much information on the history of the approved school system and its organisation today, including such aspects as intake and classifying procedures, training, formal education, after-care, staffing, administration, cost and inspection. There is also a chapter which gives a systematic account of what is done in a number of other countries. English practice earns an assessment of ' good average ' in the international scene. The second ' mainly argumentative ' half of the book is a wise, humane, canny discussion of what is to be done within the system, or within a system modified by family courts and placement in the care of the local authority. It is essentially a conservative approach from one who sees how existing institutions and relationships may be shaped effectively to better ends. Approved schools are compared with special schools for the educationally subnormal and the maladjusted ; the populations they cater for overlap, and the circumstances which determine placement in these kinds of schools are often haphazard. The factors which influence the choice of treatment are above all others the stability and adequacy of the family and the ability of parents to carry out an agreed course of action. Inadequacy in the young offender's family makes for a distinctive emphasis in treatment in an approved school : although formal educational opportunities are important, they are secondary to the need for help " with problems of adjustment to living." This implies special qualities in staff, who should be ' social workers ' as well as teachers. In order " to bring about a permanent change of a particular kind in individual behaviour " in a residential school one has to solve two problems : the needs of the individual child and the needs of the school community. These questions are considered under headings that include staff attitudes, work and leisure, links with family, and the size, location and management of approved schools.

The author appears not to hold high expectations of the contribution which psychologists can make to the understanding and treatment of youthful delinquent behaviour. This is, perhaps, not without reason—certainly the brief section on the testing of educational abilities and attainments makes depressing reading. However! In a short book he devotes a good many pages to summarising and examining descriptive typologies and forms of discipline and training He appreciates that the function of the former is to provide tools for evaluation of the effectiveness of the latter. He shares our discomfort at the paucity of evidence that approved schooling achieves its long-term aims He shows much insight into the significance of punishment, the nature of pupil-staff relationships, and the vital importance of work with the families—" aftercare begins when the boy comes in." Therefore, it is disquieting to find no reference to the work of psychologists such as Argyle, Banks, Sullivan and Grant, and Quay, to name but a few.

The approved school population is not large. It is a captive population—in the net for nearly two years on average. It is also a population about which a great deal of information is collected, by various agencies, prior to admission. Despite this, despite all that has been written on the psychological and sociological aspects of the subject, all the outpourings of conferences, despite the fact that ' juvenile delinquency ' is always news, there is distressingly little systematic evidence which Mr. Rose can cite in support of his views

C. J. PHILLIPS.

SMITH, F., and MILLER, G. A. (Edit.) (1966). *The Genesis of Language*. Cambridge, Mass.: M.I.T. Press, pp. xii+400, 80s.

In April, 1965, a conference on 'Language Development in Children' was sponsored by the Human Communication Programme of the National Institute of Child Health and Human Development, and held at Old Point Comfort, Virginia. This book is essentially a collection of the presented papers and the discussion they evoked, but unlike so many published proceedings, it has a unifying theme over and above its title. As J. J. Jenkins says in the final chapter of Reflections on the Conference: "psycholinguistics is . . . undergoing a violent and far-reaching revision." Two factors are responsible for this change in outlook the linguistic preoccupation with generative grammar and the dissatisfaction experienced by many present-day psychologists with the adequacy of reinforcement theory. The result is the underlying theme of this book which deals with the extent to which language development can be interpreted in terms of innate endowment and maturational process

It is difficult to think of a more radical revision The attempt to introduce ideas of innate equipment as the controlling source of processes which have seemed so obviously dependent on experience appears at first unnecessarily provocative, but there are good reasons for the change in approach. David McNeill, one of the principal protagonists, points out that the basic process of grammatical speech seems to be complete by the time a child is 3-and-a-half years of age and yet the speech of adults appears to be a completely random, haphazard sample, in no way contrived to instruct a child on grammar. How then can the child's knowledge of grammatical rules, his *competence*, be a product of imitative practice ? The appeal to innate 'linguistic universals' certainly offers a plausible, and more immediate, answer to the problem of the acquisition of grammar. Yet, the uneasy feeling is left in the mind of this reviewer that the *volte-face* may be premature. Though the innate argument is plausible it does not seem *so* necessary as a more thorough examination of the child's empiricism, both in regard to competence and performance, than is at present available

This book is of real interest and value to all workers in the field of education, though readers who are not already acquainted with psycholinguistics will find its own language difficult. Meta-languages cannot avoid being technical However, the effort is worthwhile, for research on the development of children's skills in their first language promises answers to a number of long-standing educational questions.

Research workers have particular cause to be grateful for the Appendix prepared by D. I. Slobin This contains abstracts of Soviet studies of child language which are not easily accessible. D J. BRUCE

STEVENSON, H. W., HESS, E. H. and RHEINGOLD, HARRIET L. (Edits.) (1967). *Early Behaviour: Comparative and Developmental Approaches*. New York· Wiley, pp. 303, 75s.

Behavioural development in infancy has recently become one of the growth areas in psychology, and here more than in most fields the contribution of animal studies to our understanding of human functioning has become evident Concepts such as critical periods, imprinting, and early stimulation have arisen mainly in the course of work with lower animals, yet have exercised a considerable influence on the gathering and interpretation of data concerning human children.

The purpose of the book under review is essentially to stress the importance of this comparative approach. It is the outcome of two conferences organised by the American Social Science Research Council, and contains nine highly polished, effectively presented research reports. Most of these are reviews of work already published: some well known, such as Fantz on visual perception in the human infant, others dealing with rather less topical and yet equally fascinating areas such as the development of handedness in cats and monkeys. Thus variety is the hallmark of this book—variety in the species studied (from doves and chicks to chimpanzees and humans), variety in research problems (such as the organisation of sucking behaviour, factors influencing choice of mate, motivational aspects of social responsiveness), and variety of methods and techniques. To some extent such a mixture makes the book all the more appealing, but in other respects a diversity is introduced which almost defeats the avowed purpose of the volume, in making it extremely difficult to go beyond the self-imposed restrictions of each chapter and discern a truly comparative trend. Few of the contributions make reference to more than one particular species or show in any other way an awareness of comparative problems. Fortunately, two chapters make good the deficiency. In one, Eibl-Eibesfeldt gives an excellent account of some ethological concepts and ranges far and wide in producing illustrative material from many points of the evolutionary scale. And in a final chapter Harriet Rheingold tackles most competently the difficult task of drawing together the different threads and enumerating those characteristics of early behaviour which appear to possess phylogenetic generality. The evolution of behaviour, it is pointed out, is still a new subject, but how our capacities to learn, to communicate, and to live socially were modified by evolutionary pressures must surely be regarded as important topics for research. There is after all, a reason for the psychologist's study of rats and mice over and above that of the convenience of dealing with subjects who cannot answer back.

H. R. SCHAFFER.

TELFORD, C. W., and SAWREY, J. M. (1967). *The Exceptional Individual.* Prentice Hall. pp. 536, no price.

There are a considerable number of books, mainly American, describing the exceptional child. This volume, by Charles W. Telford and James M. Sawrey, stands out among them. It starts with a full discussion of the basic psychological and social considerations which affect the definition and treatment of exceptionality—a concept which here includes not only the extremes of giftedness and defect but also sensorial and physical handicaps, as well as those deriving from social deviance. Many different points of view are presented in a concise manner but superficiality is avoided. A particularly useful and typical chapter in the first part of the book is given to the family and personal problems of the exceptional individual. The position is clearly stated and a substantial bibliography supports its contentions and points the student to further reading.

Subsequent sections explore different categories of handicap and deviance in turn, while at the same time stressing that " there is no one-to-one relationship between disability and handicap." While the authors' main concern throughout is for the psychological and educational aspects of their subject, it is worth noting that the physical and medical determinants of many handicaps receive serious attention. The descriptions of conditions such as mongolism, epilepsy and so on are scientific and jargon-free. The discussion of such topics as the post-natal and peri-natal causes of mental retardation and other disabilities is similarly dealt with.

The background and data of the book are, of course, American. The definitions, figures for incidence, references to school organisation and clinic provision, and much else, reflect this fact. However, so long as the reader realises this, it does not need to detract from the usefulness of the work. For one thing, many of the problems discussed are general and universal—such matters as the pros and cons of residential *v.* day treatment, the vexed question of the segregation of disabilities, the desirability or otherwise of vocational emphasis. Views on such subjects are set forth in an unbiassed manner, with supporting evidence from research findings when available Useful (but all-American) bibliographies conclude each chapter.

In a real sense the American material has, at times, unique value. For example, the author's account of 'the culturally disadvantaged' draws a picture that is particularly illuminating for the English reader, who sees his own problems writ large and differently. A particularly helpful chapter in this connection and of direct relevance, is that on the linguistic and cognitive consequences of cultural deprivation. The importance and nature of this sort of handicapping experience is extremely well conveyed.

Taken as a whole, the book is a mine of information (made accessible by adequate indexing) and each part and section provides clear and up-to-date knowledge in a scholarly but not indigestible way. It can be recommended even to more advanced students and should be on the shelves of every college of education and university education department library. O. C. SAMPSON.

NEW JOURNAL

MANPOWER AND APPLIED PSYCHOLOGY

Those who work in the field of applied psychology will be heartened by the publication of this new journal, the first of its kind to appear in Ireland and distinct from any other psychological journal in the British Isles. As the Editor, Professor P. J. R. Dempsey, of University College, Cork, explains in the first issue, the term *Manpower* in the title does not refer solely to industrial manpower, but to all human resource and effort, whether in schools, factories, fields, shops or offices. Similarly, *Applied Psychology* is used to denote the inter-relation of the various branches of psychology and the contribution of other disciplines to the systematic study of the application of human talent. It is hoped that the journal will be of interest not only to psychologists, but also to educationalists, scientists, engineers, managers and administrators in all types of employment.

The first number, which came out in March this year, is dedicated to Sir Frederic Bartlett who contributes an article on *Some Remarks about Skill.* Other contributors from this country include R. Conrad on *Training Entire Populations* and J. Annett on *Programmed Learning and the Technology of Training and Education*, while, from farther field, Professor F. J. Th. Rutten of Nijmagen University presents a short paper on *Training of Secondary School Teachers.* Among a variety of articles by, Irish authors may be mentioned a survey of Education in Agriculture by T. F. Rafferty, and an account of the Eighteenth International Congress of Psychology, in Moscow, by M. F. McHugh, in which he gathers together some of the educational implications of Western and Russian research. A novel feature, right at the end of the journal after a section on book reviews, is a couple of pages headed *Reflections*, containing a series of quotations from such varied sources as Mary Parker Follett, Ecclesiasticus and Robert Browning. This, together with a piece of light verse by D. McMahon, which appears earlier, makes a refreshing change from the heavy seriousness of some longer-established journals, while in no way detracting from the standard of the other contributions. Some of the articles are directed more to the layman than to the professional psychologist; all are noteworthy for their clarity and absence of jargon. To judge from the first issue, the new journal should have a wide circle of readers.

Manpower and Applied Psychology will appear twice yearly, in Winter and Spring, and the subscription rate is £1 per volume, 10s. per issue, in Ireland and Great Britain. Subscriptions should be sent to the Sales Manager, The Ergon Press, 45, South Mall, Cork, Ireland.

SYLVIA SHIMMIN.

PUBLICATIONS RECEIVED

The mention of a book in this list neither implies, nor precludes, a later review.

D. S. Elliott and E. T. May (1967). *Health Education.* London : Macmillan and Co., Ltd., pp. vii+115, 15s.

E. Dahlstrom (1967). *The Changing Roles of Men and Women.* London : Gerald Duckworth and Co., Ltd., pp. 298, 42s.

K. S. Holt and J. K. Reynell (1967). *Assessment of Cerebral Palsy.* London : Lloyd-Luke (Medical Books), Ltd., pp. viii+187, 40s.

T. Husén (1967). *International Study of Achievement in Mathematics—I.* London : John Wiley and Sons, Inc., pp 304.

T. Husén (1967). *International Study of Achievement in Mathematics—II.* London : John Wiley and Sons, Inc., pp. 368.

M. France (1967). *The Paradox of Guilt.* London : Hodder and Stoughton, Ltd., pp. 127, 25s.

R. Gosling, D. H. Miller, P. M. Turquet, D. Woodhouse (1967). *The Use of Small Groups in Training.* Welwyn : The Godicote Press, Ltd., pp. 144, 20s.

D. M. C. Dale (1967). *Deaf Children at Home and at School.* London : University of London Press, Ltd., pp. 272, 35s.

M. R. Heaford (1967). *Pestalozzi.* London : Methuen and Co., Ltd., pp. viii+100 22s. 6d.

H. W. Stevenson, E. H. Hess, H. L. Rheingold (1967). *Early Behaviour.*, London : John Wiley and Sons, pp. ix+303, 75s.

M. E. Franklin (1966). *Q Camp.* London : Planned Environment Therapy Trust, pp. 68, 5s.

H. C. Lindgren (1967). *Educational Psychology in the Classroom.* London : John Wiley and Sons, pp. xx+686, 68s.

B. Simon (1967). *Education : The New Perspective.* Leicester : Leicester University Press, pp 26, 3s.

C. V. Leonard (1967). *Understanding and Preventing Suicide.* Illinois : Charles C. Thomas, pp. xii+351, $11.50.

Dr. K. Lovell (1967). *Team Teaching.* Leeds : Gardham-Jowett (Printers), Ltd., pp. 52, 6s.

A. J. Peters (1967) *A Guide to the Study of British Further Education.* Slough : National Foundation for Educational Research, pp. 67, 15s.

INDEX TO VOLUME XXXVII, 1967

Articles:

Ability in Science at Ordinary Level of the General Certificate in Education: D. G. Lewis 361
Aesthetic Appreciation in Relation to Age, Variables Affecting: E. E. Rump and V. Southgate 58
Child Who Dislikes Going to School, The: S. Mitchell and M. Shepherd 32
Concepts of Time in Secondary School Children of Above Average I.Q.: K. W. Rodgers 99
Developmental Changes in Children's Drawings: M. P. M. Richards and H. E. Ross 73
Discrimination Learning and Discrimination Learning Sets in Subnormal Children: E. A. Lunzer and I. Hulme 175
East African Pupils, Administration of Group Intelligence Tests to: P. E. Vernon 282
Effects of Omitting Branches and Questions from a Scrambled Text: K. D. Duncan and T. Gilbert 314
Effects of the 'Eleven-Plus' Result on the Subsequent Careers of Pupils, The: E. D. Cave 41
English Proficiency of Overseas Students, The: A. Davies 165
Environmental Influences on Responses to Questionnaires: M. D. Shipman 54
Evaluations of Three Factors that Might Influence Responses to Tests of Academic Achievement Motivation: C. E. Buxton 241
Failures in Children to Sort by Two Different Dimensions on Successive Trials, The Causes of: P. E. Bryant 320
Family Environment and 11+ Success: Some Basic Predictors: D. F. Swift 10
Fluid and Crystallized General Intelligence Checked at the 5/6-year-old Level, The Theory of: R. B. Cattell 209
Gifted Child, Some Aspects of a Study of the: K. Lovell and J. B. Shields 201
Influence of Reading Methods on Spelling, The: M. L. Peters 47
Intelligence and Family Size, 1949-1965: J. D. Nisbet and N. J. Entwistle 188
Intelligence Assessment: A Theoretical and Experimental Approach: H. J. Eysenck 81
Multiple-Choice *versus* Open-Ended Answering in a Vocabulary Test, An Experiment on: A. W. Heim and K. P. Watts 339
Personality Pattern of Creative Artists, The: P. G. Cross, R. B. Cattell and H. J. Butcher 292
Premature Retirement of Women Teachers from Girls' and Mixed Secondary Schools: R. R. Dale 329
Productive Thinking in Retarded and Non-Retarded Children: J. F. Cawley and D. V. Chase 356
Road Safety Education for Young Children: K. Pease and B. Preston 305
Selection for Remedial Education, Follow-up of an Experiment in: H. Lytton 1
Self-Concept in the Psychological Guidance of Adolescents, The: H. A. S. Zahran 225
Streaming and Season of Birth in Secondary Schools, Correlation between: P. Sutton 300
Teachers' Ratings of Schoolboys' Behaviour Related to Patterns of Scores on New Junior Maudsley Inventory: H. B. Gibson 347
Teacher's Role, The: A Comparative Study of the Conceptions of College of Education Students and Head Teachers: D. S. Finlayson and L. Cohen 22
Value Patterns of Engineering, Law, Medical and Teacher-Training Students in India: S. K. Pal 371
WISC in a Psychiatric Group, The Stability of the: R. K. Turner, A. Mathews and S. Rachman 194

Research Notes:

Aural Analysis of Sounds, The: C. Carver 379
Birth Order, Achievement-Motivation and Academic Attainment: F. H. Farley 256
Co-education as a Possible Influence on the Desire for Marriage: R. R. Dale 119
Comments on "An Investigation of the Craig Reader Course," by K. D. Duncan and R. Foster: G. H. McLaughlin 260
Effect of Restriction of Range Upon Correlation Coefficients, The: S. Wiseman 248
Intelligence Tests, Advanced Level and First-Year University Examination Marks for Predicting Degree Classification, The Relative Value of two High Level: G. W. Pilkington and G. J. Harrison 382
Junior Maudsley Personality Inventory, Some Evidence on the Validity of the: A. H. Mehryar 375
Minnesota Tests of Creative Thinking—a Note on Scorer Reliability in Follow-up Studies with English Primary School Children, The: R. J. Goldman and D. F. Clarke 115

Note on the Article " Differences in Test Scores and in the Gaining of Selective Places for Scottish Children and those in England and Wales," by Douglas, Ross, Maxwell and Walker : H. Lytton .. 124
Oblique Factorization of Wallach and Kogan's " Creativity " Correlations : J. Ward 380
Performance of American Pupils on the Coloured Progressive Matrices, The : Z. M. Cantwell .. 389
Predicting Reading Ages on Neale's Analysis of Reading Ability : W. Yule .. 252
Progressive Matrices (1938) and the Dominoes (D48) Tests : A Cross-Cultural Study, The : A. Abi Rafi .. 17
Regression Analysis Programmes for the Identification of Important Variables in Educational and Psychological Experiments . P. Williams .. 112
Reliability and Validity of the Goodenough-Harris Drawing Test, The : W. Yule, L. Lockyer and A. Noone .. 110
Valentine's Reasoning Tests in an Investigation into Transfer from Mathematics to Reasoning, The Use of : J. K. Backhouse .. 121
Wastage of Ability Amongst the Intellectually Gifted : H. G. Armstrong .. 257

Critical Notices :

Examining at 16+. The Schools Council (1966) : N. France .. 406
Gifted Children, by M. Branch and A. Cash : C. Burt .. 143
i.t.a. Symposium, The : G. Keir .. 408
The Public Schools : A Factual Survey, by Graham Kalton : R. R. Dale .. 149

Thesis Abstracts :

Attainments in Reading and Spelling of Children Who Learned to Read Through the Initial Teaching Alphabet, The : T. D. Swales .. 126
Composition Performance of Ten-year-old Children, Factors Affecting the : D. Sharples .. 137
Effects of Increasing the Teachers' Knowledge of their Pupils' Self Pictures, Some : J. A. Chadwick .. 129
Idea of God among Roman Catholic Children, aged Seven and Eight, The : D. Martyn 399
Motivation, Personality, Intelligence and School Attainment in a Secondary Modern School, The Relationship Between : M. E. Ainsworth .. 135
Objective Indices in the Assessment of Essays : R. A. Newbery .. 403
Objective Tests in English, The Validity of : F. H. Sparrow .. 393
Opinions on Coeducation of a Group of Secondary Modern Children in the Manchester Area, their Teachers and their Parents, The : A. Nicholls .. 396
Personality Attributes of Stammering, Enuretic and School-phobic Children : A. McHale .. 400
Personality Characteristics and Attitude to School : G. Regan .. 127
Personality Factors and Attitudes of Mature Training College Students : E. Solomon 140
Personality Measures Associated with Over and Under Achievement in English and Arithmetic, An Investigation of . L. W. Ridding .. 397
Psychological Factors Associated with the Sociometric Status of Children Attending a Comprehensive School in Breconshire : H. H. Richards .. 261
Reading Efficiency of Students at a Technical Teacher Training Centre, An Investigation into : R. L. James .. 391
Religious Attitude of Secondary Modern County School Pupils : W. P. L. Johnson .. 132
Vertical Grouping and Horizontal Grouping in the Infant School, A Comparison of : M. A. Mycock .. 133

Authors :

Ainsworth, M. E. : The Relationship Between Motivation, Personality, Intelligence and School Attainment in a Secondary Modern School .. 135
Armstrong, H. G. : Wastage of Ability Amongst the Intellectually Gifted .. 257
Backhouse, J. K. : The Use of Valentine's Reasoning Tests in an Investigation into Transfer from Mathematics to Reasoning .. 121
Bryant, P. E. : The Causes of Failures in Children to Sort by Two Different Dimensions on Successive Trials .. 320
Burt, C. : Gifted Children, by M. Branch and A. Cash .. 143
Buxton, C. E. : Evaluations of Three Factors that Might Influence Responses to Tests of Academic Achievement Motivation .. 241
Cantwell, Z. M. . The Performance of American Pupils on the Coloured Progressive Matrices .. 389
Carver, C. : The Aural Analysis of Sounds .. 379
Cattell, R. B. : The Theory of Fluid and Crystallized General Intelligence Checked at the 5/6-year-old Level .. 209

Cave, E. D.: The Effects of the 'Eleven-Plus' Result on the Subsequent Careers of Pupils .. 41
Cawley, J. F., and Chase, D. V.: Productive Thinking in Retarded and Non-Retarded Children .. 356
Chadwick, J. A.: Some Effects of Increasing the Teachers' Knowledge of their Pupils' Self Pictures .. 129
Cross, P. G., Cattell, R. B., and Butcher, H. J.: The Personality Pattern of Creative Artists .. 292
Dale, R. R.: Co-education as a Possible Influence on the Desire for Marriage .. 119
Dale, R. R.: The Public Schools: A Factual Survey, by Graham Kalton .. 149
Dale, R. R.: Premature Retirement of Women Teachers from Girls' and Mixed Secondary Schools .. 329
Davies, A.: The English Proficiency of Overseas Students .. 165
Duncan, K. D., and Gilbert, T.: Effects of Omitting Branches and Questions from a Scrambled Text .. 314
Eysenck, H. J.: Intelligence Assessment: A Theoretical and Experimental Approach .. 81
Farley, F. H.: Birth Order, Achievement-Motivation and Academic Attainment .. 256
Finlayson, D. S., and Cohen, L.: The Teacher's Role: A Comprehensive Study of the Conceptions of College of Education Students and Head Teachers .. 22
France, N.: The Schools Council (1966). Examining at 16+. The Report of the Joint G.C.E./C.S.E. Committee of the Schools Council .. 406
Gibson, H. B.: Teachers' Ratings of Schoolboys' Behaviour Related to Patterns of Scores on the New Junior Maudsley Inventory .. 347
Goldman, R. J., and Clarke, D. F.: The Minnesota Tests of Creative Thinking—A Note on Scorer Reliability in Follow-up Studies with English Primary School Children .. 115
Heim, A. W. and Watts, K. P.: An Experiment on Multiple-Choice *versus* Open-Ended Answering in a Vocabulary Test .. 339
James, R. L.: An Investigation into the Reading Efficiency of Students at a Technical Teacher Training College .. 391
Johnson, W. P. L.: The Religious Attitude of Secondary Modern County School Pupils .. 132
Keir, G.: The i.t.a. Symposium .. 408
Lewis, D. G.: Ability in Science at Ordinary Level of the General Certificate of Education .. 361
Lovell, K., and Shields, J. B.: Some Aspects of a Study of the Gifted Child .. 201
Lunzer, E. A. and Hulme, I.: Discrimination Learning and Discrimination Learning Sets in Subnormal Children .. 175
Lytton, H.: Follow-up of an Experiment in Selection for Remedial Education .. 1
Lytton, H.: Note on the Article "Differences in Test Scores and in the Gaining of Selective Places for Scottish Children and those in England and Wales," by Douglas, Ross, Maxwell and Walker .. 124
McHale, A.: An Investigation of Personality Attributes of Stammering, Enuretic and School-phobic Children .. 400
McLaughlin, G. H.: Comments on "An Investigation of the Craig Reader Course," by K. D. Duncan and R. Foster .. 260
Martyn, D.: The Idea of God Among Roman Catholic Children Aged Seven and Eight .. 399
Mehryar, A. H.: Some Evidence on the Validity of the Junior Maudsley Personality Inventory .. 375
Mitchell, S., and Shepherd, M.: The Child Who Dislikes Going to School .. 32
Mycock, M. A.: A Comparison of Vertical Grouping and Horizontal Grouping in the Infant School .. 133
Newbery, R. A.: Objective Indices in the Assessment of Essays .. 403
Nicholls, A.: The Opinions on Co-education of a Group of Secondary Modern Children in the Manchester Area, Their Teachers and Their Parents .. 396
Nisbet, J. D. and Entwistle, N. J.: Intelligence and Family Size, 1949-1965 .. 188
Pai, S. K.: Value Patterns of Engineering, Law, Medical and Teacher-Training Students in India .. 371
Pease, K., and Preston, B.: Road Safety Education for Young Children .. 305
Peters, N. L.: The Influence of Reading Methods on Spelling .. 47
Pilkington, G. W., and Harrison, G. J.: The Relative Value of Two High Level Intelligence Tests, Advanced Level, and First-Year University Examination Marks for Predicting Degree Classification .. 382
Rafi, A. Abi: The Progressive Matrices (1938) and the Dominoes (D48) Tests: A Cross-Cultural Study .. 117
Regan, G.: Personality Characteristics and Attitude to School .. 127

Richards, H. I.: Psychological Factors Associated with the Sociometric Status of Children Attending a Comprehensive School in Breconshire 261
Richards, M. P. M., and Ross, H. E.: Developmental Changes in Children's Drawings 73
Ridding, L. W.: An Investigation of Personality Measures Associated with Over and Under Achievement in English and Arithmetic 397
Rogers, K. W.: Concepts of Time in Secondary School Children of Above Average I.Q. 99
Rump, E. E., and Southgate, V.: Variables Affecting Aesthetic Appreciation in Relation to Age 58
Sharpes, D.: Factors Affecting the Composition Performance of Ten-year-old Children 137
Shipman, M. D.: Environmental Influences on Response to Questionnaires 54
Solomon, E.: Personality Factors and Attitudes of Mature Training College Students 140
Sparrow, F. H.: The Validity of Objective Tests in English 393
Sutton, P.: Correlation between Streaming and Season of Birth in Secondary Schools 300
Swales, T. D.: The Attainments in Reading and Spelling of Children Who Learned to Read Through the Initial Teaching Alphabet 126
Swift, D. F.: Family Environment and 11+ Success: Some Basic Predictors 10
Turner, R. K., Mathews, A., and Rachman, S.: The Stability of the WISC in a Psychiatric Group 194
Vernon, P. E.: Administration of Group Intelligence Tests to East African Pupils 282
Ward, J.: An Oblique Factorization of Wallach and Kogan's "Creativity" Correlations 380
Williams, P.: Regression Analysis Programmes for the Identification of Important Variables in Educational and Psychological Experiments 112
Wiseman, S.: The Effect of Restriction of Range Upon Correlation Coefficients 248
Yule, W.: Predicting Reading Ages on Neale's Analysis of Reading Ability 252
Yule, W., Lockyer, L., and Noone, A.: The Reliability and Validity of the Goodenough-Harris Drawing Test 110
Zahran, H. A. S.: The Self-Concept in the Psychological Guidance of Adolescents. 225

NOTES TO CONTRIBUTORS

1.—The Editor is glad to consider articles in the fields of educational and child psychology, whether written from the theoretical, clinical, experimental or psychometric standpoints. It is, however, desirable that statistical matter (formulae, tables, discussions of techniques, etc.) should be kept to a minimum, so that the *Journal* will be of interest to statistically untrained readers. Such concepts as the mean, percentiles, standard deviation, correlation, chi-squared, standard error, critical ratio, significance and reliability, and general or group factors, can be employed in the text. But more complex concepts and methods should be explained or confined to footnotes or an appendix, or published elsewhere.

Articles which have been, or are to be, published elsewhere cannot be accepted.

2.—*Length.* The usual range of length is from 2,500 to 4,500 words, and only very exceptionally will articles of more than 6,000 words be published. Short research reports of 500—1,000 words, such as outlines of theses, are welcomed. These should normally be submitted through the Head of the Department in which the research was carried out.

3.—*Arrangement.* Articles should be *preceded* by a summary or abstract of between 100 and 300 words, embodying the main conclusions. This should be typed directly after the title and author's name and place of work. A contents analysis, or list of headings, need *not* be included, but successive sections of the article should receive numbered headings, e.g. :

I. INTRODUCTION.
II. OUTLINE OF PREVIOUS RESEARCH.
III. PLAN OF THE INVESTIGATION.
  (i) *Subjects and materials.*
  (ii) *Experimental Procedure.*
IV. STATEMENT OF RESULTS.
V. DISCUSSION AND CONCLUSIONS.
VI. REFERENCES.

4.—*Typing.* All papers must be typed (double or 1½-spaced) on one side of quarto or foolscap paper. There should be margins of at least one inch all round. Authors should consult articles in current numbers of the *Journal*, and copy their conventions in respect of capitalising the title and main section headings, italicising author's place of work and sub-section headings, indentation of paragraphs, the form of references and tables, etc.

5.—*Footnotes* should be few in number, and should be indicated in the text by using the signs *, †, ‡, §.

6.—*References* should normally be indicated by date following the author's name in the text, e.g., Spearman (1904), or (Spearman, 1904, p 160), or—where there are two or more publications in the same year—Burt (1939a). However, it is occasionally more convenient to number the references throughout, and to type numbers in the text above the line, e.g., Spearman.^{4} When there are two or three references only, these may be placed in footnotes in the text, instead of in a separate section at the end.

The list of references at the end should always be in alphabetical order of author's names, and—for any one author—by order of date. The following standard form of references should be observed :

DEARBORN, W. F., and ROTHNEY, J. W. M. (1941). *Predicting the Child's Development.* Cambridge, Mass. : Sci.-Art.

(Note that Ltd., etc., after publisher's names, also the numbers of pages in the book, can be omitted.)
BURT, C. (1939a). The relations of educational abilities. *Brit. J. Educ. Psychol.*, 9, 55-71

(Note the absence of inverted commas and of capital initials in the titles of articles. The number of the volume is in arabic, not roman numerals. The part of the journal is not given, but page numbers are included in full.)

The World List of abbreviations (as given in the December issues of *Psychological Abstracts*) is used for journals.

7.—*Tables and figures* are costly to print, and should be used only when they save space and/or are essential for effective presentation. The same material should not be presented in a table *and* a graph. Each table or figure should be numbered (in arabic figures), on a separate sheet, and its approximate position indicated in the typescript. Drawings should be in Indian ink on heavy, unruled paper, the lettering on graphs being in pencil. Lines in tables should be ruled in pencil. Each table or graph should have a short title, sufficient for understanding without searching in the text

8.—*Corrections.* The Editor cannot consider for publication articles which are seriously deficient in presentation. The typescript submitted (after preliminary consultation, if desired) should represent the final form in which the author wishes the paper to appear. Since changes in proof are costly, out of all proportion to the original setting, authors must be charged for any changes, insertions or deletions other than printer's errors).

9.—*Reprints.* Authors receive fifty copies of their papers free. Extra copies, in multiples of fifty, may be had at cost price if the order is given when the proof is returned.

COMMUNICATION AND COMPREHENSION

S. J. M. Trenaman

This volume records the only large-scale British experimental study yet to be undertaken on the subject of television as a means of communication. The recent extensive development of television as an educational medium has set a premium upon balanced statistical information in this field. 35*s*

STUDIES IN CHILD DEVELOPMENT

Published by Longmans in association with the National Bureau for Co-operation in Child Care.

THE COMMUNITY'S CHILDREN

Jessie Parfit

A comprehensive analysis of the child care services. The book gives the routine facts and shows the emotional problems and special difficulties which face professional staff, parents and children. 15*s*

FOSTER HOME CARE—FACTS AND FALLACIES

R. Dinnage and M. L. Kellmer Pringle

A review of research in the United States, Canada and Great Britain between 1948 and 1965. 30*s*

RESIDENTIAL CHILD CARE—FACTS AND FALLACIES

R. Dinnage and M. L. Kellmer Pringle

A review of research in the United States, Canada, Great Britain, Belgium, France, Israel and West Germany between 1948 and 1965. 30*s*

Already published :

FAMILY ADVICE SERVICES:

An exploratory study of a sample of such services organised by Children's Departments in England.

Aryeh Leissner 12*s* 6*d*

FOUR YEARS ON

S. Gooch and M. L. Kellmer Pringle 17*s* 6*d*

11,000 SEVEN-YEAR-OLDS

M. L. Kellmer Pringle, N. R. Butler and R. Davie 21*s*

ADOPTION—FACTS AND FALLACIES

M. L. Kellmer Pringle 21*s*

1724

Longmans

The British Journal of

PSYCHOLOGY

Edited by ARTHUR SUMMERFIELD

Assisted by

JAMES HOGG and PETER MITTLER

This Journal which is published for the British Psychological Society is particularly concerned with experimental psychology and methods of testing for, and measuring psychological variants. The large book review section which covers publications in German as well as English is important.

The subscription price is £4 10s. per volume, and 30s. per single issue. Published quarterly, there are 4 parts to a volume.

Orders may be placed with a bookseller, or sent direct to the publishers.

CAMBRIDGE UNIVERSITY PRESS

Bentley House,
200 Euston Road, London, N.W.1

The British Journal of

Social and Clinical Psychology

Contents of Vol. 6, Part 3, September, 1967

MORRISON, A. and McINTYRE, D. Changes in opinions about education during the first year of teaching.
STRONGMAN, K. T. and WOOSLEY, J. Stereotyped reactions to regional accents.
GIBSON, H. B. Self-reported delinquency among schoolboys, and their attitudes to the police.
WALTERS, R. H. and AMOROSO, D. M. Cognitive and emotional determinants of the occurrence of imitative behaviour.
STRINGER, P. Masculinity-femininity as a possible factor underlying the personality responses of male and female art students.
LINDSAY, J. S. B. The structure within groups.
GERVER, D. Linguistic rules and the perception and recall of speech by schizophrenic patients.
BRYANT, P. E. Verbalization and immediate memory of complex stimuli in normal and severely subnormal children.
MAIR, J. M. M. A comparison of two grid forms.
DAVIES, A. D. M. Age and the Memory-for-Designs test
BOOK REVIEWS.

The subscription price per volume of four parts is 90s. net. Orders may be sent to your bookseller or direct to the publishers

CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200, Euston Road, London, N.W.1

THE TEACHING REVOLUTION

W. Kenneth Richmond

The structure, methods and approach of teaching are undergoing rapid change. This stimulating book draws together many separate strands of development, including programmed learning, closed-circuit television and curriculum reform projects, to present an overall view of present-day changes, how they are being met, and what pattern for the future is emerging.

With 12 photographs and 11 diagrams 36s

AN INTRODUCTION TO EDUCATIONAL PSYCHOLOGY

E. Stones

" Here, at last, is the book we have been waiting for, an introduction to educational psychology which links theory to practice, and which draws on the most advanced thinking and recent scientific developments." *Forum* 36s

University Paperback 18s

EXPERIMENT AND TRADITION IN PRIMARY SCHOOLS

D. E. M. Gardner

The result of twenty-five years' pioneering work in educational research, which compares the attitudes and attainments of children whose primary schooling has been entirely ' experimental ' with those of children from ' traditional ' primary schools. 22s 6d

The Library of Educational Thought

JOHN LOCKE: Prophet of Common Sense 25s

M. V. C. Jeffreys *Education Paperback 12s 6d*

ARISTOTLE 25s

C. Winn *and* **M. Jacks** *Education Paperback* 12s 6d

PESTALOZZI 25s

Michael Heafford *Education Paperback* 12s 6d

GEORG KERSCHENSTEINER 25s

Diane Simons *Education Paperback* 12s 6d

METHUEN

Perceptual Development in Children

Edited by Aline H. Kidd and Jeanne L. Rivoire

Perceptual Development in Children is a compilation of articles by twenty-five authors, two of them—M. D. Vernon and Kenneth Lovell—British, and the rest American. The book is divided into six major categories : Psychological Factors in the Development of Perception, Theories of Perception. Visual and Auditory Perception, Social Aspects of Perception, Affective Aspects of Perception, and Cognitive Aspects of Perception. *Perceptual Development in Children* will be of interest to all departments of psychology and education and should be especially helpful because it is the only compendium published in Britain which gives the results of recent American research on the subjects chosen. Originally published in the U.S.A. by International Universities Press Inc. .

Just Published 100/- net

Educational Psychology and Children

K. LOVELL, B.SC., M.A., PH.D.

Ninth Edition

A comprehensive and up-to-date textbook of educational psychology. This latest edition, the ninth since its original publication in 1958, has been completely revised, fresh material incorporated and all references brought up-to-date.

18/- net

FROM ANY BOOKSELLER

UNIVERSITY OF LONDON PRESS LTD St. Paul's House, Warwick Lane, London, EC4

A SYSTEM OF COMPUTER PROGRAMS
FOR USE IN PSYCHOLOGY AND EDUCATION

by

H. J. HALLWORTH

assisted by

ANN BREBNER

This handbook explains the preparation of data for a system of computer programs which will perform a variety of statistical computations frequently used in the social sciences.

The separate programs were originally developed in the Education Department at the University of Birmingham and are published by the British Psychological Society on the recommendation of a committee appointed to consider the development of computer programming. The programs have been organised into a system which may be used by those with little or no previous experience of electronic computers. It is merely assumed that the user is acquainted with a teleprinter or a card punching machine.

Price 10/- (post free)

Available from :

THE BRITISH PSYCHOLOGICAL SOCIETY
TAVISTOCK HOUSE SOUTH
TAVISTOCK SQUARE
LONDON, W.C.1

SOCIAL SCIENCE
PAPERBACKS

Influences on Parent Behavior

LOIS MEEK STOLZ

Data from over three hundred interviews provide the basis of this inquiry into the past and present influences that have affected the child-rearing practices of a sample of American parents. The findings have implications both for future research and for parent education.
Stanford Studies in Psychology November
65s net

Men in Middle Life

KENNETH SODDY
WITH MARY C. KIDSON

An interdisciplinary, cross-cultural study of the physical, mental, and social welfare of middle-aged men in the context of their family, community, and occupational relationships.
Cross-cultural Studies in Mental Health, 3
November 63s net

Neurosis in the Ordinary Family

A PSYCHIATRIC SURVEY

ANTHONY RYLE

'So much redundant activity and writing is justified by the title of "research" that it is an unusual pleasure to come across a really worthwhile project, clearly described... a most important and stimulating book.' Hugh Freeman, *New Society*
Mind and Medicine Monographs No. 15
32s 6d net

The Family and Individual Development

D. W. WINNICOTT

By the author of *The Child and the Family* and *The Child and the Outside World*

'These essays are packed with the author's very personal wisdom, humanity, and charm, and will illuminate the understanding of any adult who reads them. *Archives of Disease in Childhood* **30s net**

A Textbook of Health Education

DENIS PIRRIE AND
A. J. DALZELL-WARD

SSP 3 **18s net**
Hardback 35s net

Social Psychology

W. J. H. SPROTT

SSP 4 (Methuen) **18s net**
Hardback 35s net

The Unconscious Before Freud

LANCELOT LAW WHYTE

SSP 19 **17s 6d net**
Hardback 28s 6d net

Nurse and Patient

G. BURTON

SSP 22 **November** **17s 6d net**
Hardback 21s net

A Study of Brief Psychotherapy

D. H. MALAN

SSP 23 **November** **21s net**
Hardback 35s net

Community as Doctor

NEW PERSPECTIVES ON A THERAPEUTIC COMMUNITY

ROBERT N. RAPOPORT

SSP 24 **24s net**
Hardback 45s net

The Education of Technologists

AN EXPLORATORY CASE STUDY AT BRUNEL

MARIE JAHODA

SSP 25 **November** **17s 6d net**
Hardback 30s net

Psychology and Social Problems

MICHAEL ARGYLE

SSP 26 (Methuen) **16s net**

TAVISTOCK PUBLICATIONS

THE BRITISH JOURNAL OF MEDICAL PSYCHOLOGY

Edited by Dr. THOMAS FREEMAN

Volume 40, *Part* 3. CONTENTS *September*, 1967

D. L. Rubinfine. Notes on a Theory of Reconstruction.
A. C. Woodmansey. Emotions and The Motions: An Enquiry into the Causes of Functional Disorders of Defaecation.
W. S. Inman. Emotion, Cancer and Time.
H. L. Martin. The Significance of Discussion with Patients about Their Diagnosis and Its Implications.
Georges Reding, Lois A. Charles and Michael B. Hoffman. Treatment of The Couple by a Couple ; II Conceptual Framework, Case Presentation and Follow Up Study.
Phillida Salmon, Judith Bramley and A. Presley. The Word-in-Context Test as a measure of Conceptualisation in Schizophrenics with and without Thought Disorder.
I. M. Marks. Components and Correlates of Psychiatric Questionnaires.
E. V. Leckie and R. J. F. Withers. A Test of Liability to Depressive Illness.
Eugene Ziskind. Sociogenic Models for Mental Disease.
A. W. Kushner. Two cases of Auto-Castration.

Subscription £4 10*s. per volume of* 4 *parts. Single parts* 30*s. plus postage*

Orders may be placed with any bookseller, or direct to :
CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200 Euston Road, London, N.W.1

AUSTRALIAN JOURNAL OF PSYCHOLOGY

Editor : Dr. RONALD TAFT

Vol. 19, *No.* 1 *April*, 1967

N. T. Feather. Evaluation of Religious and Neutral Arguments in Religious and Atheist Student Groups.
J. R. Trotter. The Psychophysics of Melodic Interval : Definitions, Techniques, Theory and Problems.
I. D. John. A Statistical Decision Theory of Simple Reaction Time.
I. D. John. Sequential Effects of Stimulus Characteristics in a Serial Reaction Time Task.
Elizabeth M. Arnold. Is Temporal Integration a Distinct Mental Ability ?
J. S. Baily and G. Singer. Behavioural Compensation Through Information Feedback and Transformed Visual Input.
Peter J. Livesey. The Hebb-Williams Elevated Pathway Test : A Comparative Study of Rat, Rabbit and Cat Performance.
J. K. Collins. Freedom and Restriction of the Limbs in the Kinesthetic Spatial Aftereffect.
G. D. Wilson and R. A. M. Gregson. Effects of Illumination on Perceived Intensity of Acid Tastes.
M. Miller and R. A. Champion. Amount of Food Available on Entry into the Goal Box as a Variable in Instrumental Learning.

Published by the Australian Psychological Society
Registered at the G.P.O., Melbourne, for transmission by post as a periodical

The Australian Journal of Psychology is issued by the Australian Psychological Society three times a year, in April, August and December.
The Annual subscription for members of the British Psychological Society is 27s. (sterling) payable to the Honorary Treasurer of the B.P.S., Tavistock House South, Tavistock Square, London, W.C.1. A special rate of $U.S.5.00 is available to members of the American Psychological Association. For other non-members, rates are $A5.50, 50s. (U.K.) and $U.S.7·50 in U.S.A.

ROUTLEDGE & KEGAN PAUL

Now available in paperback

A TEACHER'S GUIDE TO READING PIAGET

Molly Brearley and Elizabeth Hitchfield

Respectively, Principal and Senior Lecturer, Froebel Educational Institute

'Deserves to succeed in its object of giving teachers an *aperitif* prior to reading Piaget in the raw. The authors have confined themselves to looking at eight specific aspects of Piaget's work: number, measurements, knots, perspective, co-ordinates, floating and sinking, moral judgment, and the behaviour of babies, and they wisely adopt the policy of describing an experiment prior to discussing its theoretical implications. What the writers have to say will be particularly useful to teachers in nursery and primary schools.'—*Teachers World* **8s 6d**

Library edition still available 18s

Scottish Standardisation of the Wechsler Intelligence Scale for Children

The 55th Publication of the
SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION

The W.I.S.C. has become one of the most widely used and respected means of training intelligence in the English-speaking world, but as its findings were originally based on conditions in the U.S.A. further research was necessary to make it entirely suitable for use in Scotland.

The S.C.R.E. administered over two thousand individual tests to a control group of children of the same age in order to establish norms of Scottish intelligence comparable to American ones. The latest publication is a report on this research with comments on the various sub-tests used, the scaling of the scores, sex differences and answer patterns.

11/6 net

From any bookseller

UNIVERSITY OF LONDON PRESS LTD

St. Paul's House, Warwick Lane, London, E.C.4

OCCUPATIONAL PSYCHOLOGY

Editor : Alec Rodger

APRIL and JULY, 1967 VOLUME 41, Nos. 2 and 3

CONTENTS

Putting Psychology to Work. By A. B. Cherns
Social Factors in Satisfaction with Swiftly Rotating Shifts. By A. I. I. Wedderburn.
Industrial Training : Current Research and Future Needs. By Alec Martin.
Ministry of Labour Occupational Guidance Units. By John Crinnion.
Occupational Difficulties and Distastes among Managers. By J. D. Pincus.
Labour Turnover in Two Public Service Organisations. By Sylvia Downs.
A Study in Adult Retraining : the Gas Man. By Bernard Dodd.

Annual Subscription 40 *shillings, overseas* 50 *shillings*

National Institute of Industrial Psychology
14, Welbeck Street, London, W.1.

The Birmingham Printers Ltd

PRINTERS AND BOOKBINDERS

High Class
Technical Publications

42-44 HILL STREET, BIRMINGHAM 5

Telephone : MIDLAND 0196-7

OTHER PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology.
The British Journal of Psychology (Monograph Supplements).
The British Journal of Medical Psychology.
The British Journal of Mathematical and Statistical Psychology.
The British Journal of Social and Clinical Psychology.

The British Journal of Psychology and *The British Journal of Medical Psychology* appear in quarterly parts in volumes of about 350 pages, subscription price of each, payable in advance, 60s. per volume (post free). Subscriptions may be sent to any Bookseller or to the Cambridge University Press, Bentley House, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $10.00. Enquiries should be addressed to the Cambridge University Press, American Branch, 32, East 57th Street, New York 22.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November, in volumes of about 184 pages. The subscription rate payable in advance is 30s. 6d. per volume (post free). The subscription price in the U.S.A. is $4.50 (post free). Subscriptions should be sent to the Society's Agents, Taylor and Francis, Ltd., Red Lion Court, Fleet Street, E.C.4, or to any Bookseller.

The British Journal of Social and Clinical Psychology is issued three times a year in February, June and November, in volumes of approximately 240 pages. The subscription price of each, payable in advance, is 50s. (post free). Subscriptions may be sent to any Bookseller, or to the Cambridge University Press, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $8.50. Enquiries should be sent to the Cambridge University Press, American Branch, 32, East 57th Street, New York 22. Separate parts will be sold at 20s.: $3.50 in the United States.

A series of *Monograph Supplements* is issued in connection with the *British Journal of Psychology*, but is not included in the subscription for that journal. Orders may be sent to any Bookseller or to the Cambridge University Press, 200, Euston Road, London, N.W.1, from whom details of these publications may be obtained.

MANUSCRIPTS

For the *British Journal of Psychology* should be sent to Professor ARTHUR SUMMERFIELD, Department of Psychology, Birkbeck College, Malet Street, London, W.C.1.

For the *British Journal of Medical Psychology*, should be sent to Dr. THOMAS FREEMAN, Lansdowne Clinic, 4, Royal Crescent, Kelvingrove, Glasgow, C.3.

For the *British Journal of Educational Psychology*, should be sent to Professor J. D. NISBET, King's College, Aberdeen.

For the *British Journal of Mathematical and Statistical Psychology*, should be sent to Dr. R. J. AUDLEY, Department of Psychology, University College, Gower Street, London, W.C.1.

For the *British Journal of Social and Clinical Psychology*, should be sent to Mr. MICHAEL ARGYLE (Social Psychology Editor), Institute of Experimental Psychology, 1, South Parks Road, Oxford;

or to:

Dr. RALPH HETHERINGTON (Department of Studies in Psychological Medicine), 6, Abercromby Square, Liverpool, 7.

For the *Monograph Supplements*, to Professor IAN M. L. HUNTER, Department of Psychology, University of Keele, Keele, Staffs., whose decision regarding the publication of all *Monograph Supplements* will be final. No MS. can be considered for publication in any year unless it is in the hands of the Editor by 1st September of the preceding year.

OFF-PRINTS

Contributors receive 50 free copies of their article and may buy more in units of 50.

Members of the British Psychological Society receive all the Society's Journals at reduced rates.

Information on membership and further activities may be obtained from the Secretary, Tavistock House South, Tavistock Square, London, W.C.1.

PUBLISHED BY THE ASSOCIATION OF TEACHERS IN COLLEGES AND DEPARTMENTS OF EDUCATION

Education for Teaching is published three times a year at mid-term in November, February and May, and distributed free to members of the Association. To non-members it costs 5s. per issue, post free (subscription 14s. per annum).

Members of the A.T.C.D.E. receive the *British Journal of Educational Psychology* at a reduced rate.

Information on membership and further activities may be obtained from the Honorary Secretary, 151, Gower Street, London, W.C.1.

CONTENTS

P. E. VERNON: Administration of Group Intelligence Tests to East African Pupils 282
P. G. CROSS, R. B. CATTELL and H. J. BUTCHER: The Personality Pattern of Creative Artists . . . 292
P. SUTTON: Correlation between Streaming and Season of Birth in Secondary Schools 300
K. PEASE and B. PRESTON: Road Safety Education for Young Children . . . 305
K. D. DUNCAN and T. GILBERT: Effects of Omitting Branches and Questions from a Scrambled Text 314
P. E. BRYANT: The Causes of Failures in Children to Sort by Two Different Dimensions on Successive Trials 320
R. R. DALE: Premature Retirement of Women Teachers from Girls' and Mixed Secondary Schools 329
A. W. HEIM and K. P. WATTS: An Experiment on Multiple-Choice *versus* Open-Ended Answering in a Vocabulary Test. 339
H. B. GIBSON: Teachers' Ratings of Schoolboys' Behaviour Related to Patterns of Scores on the New Junior Maudsley Inventory 347
J. F. CAWLEY and D. V. CHASE: Productive Thinking in Retarded and Non-Retarded Children 356
D. G. LEWIS: Ability in Science at Ordinary Level of the General Certificate of Education 361
S. K. PAL: Value Patterns of Engineering, Law, Medical and Teacher-Training Students in India 371

RESEARCH NOTES:

A. H. MEHRYAR: Some Evidence on the Validity of the Junior Maudsley Personality Inventory 375
C. CARVER: The Aural Analysis of Sounds 379
J. WARD: An Oblique Factorization of Wallach and Kogan's "Creativity" Correlations 380
G. W. PILKINGTON and G. J. HARRISON. The Relative Value of Two High Level Intelligence Tests, Advanced Level, and First-Year University Examination Marks for Predicting Degree Classification 382
Z. M. CANTWELL: The Performance of American Pupils on the Coloured Progressive Matrices 389

THESES ABSTRACTS:

R. L. JAMES: An Investigation into the Reading Efficiency of Students at a Technical Teacher Training College 391
F. H. SPARROW: The Validity of Objective Tests in English 393
A. NICHOLLS: The Opinions on Co-education of a Group of Secondary Modern Children in the Manchester Area, Their Teachers and Their Parents 396
L. W. RIDDING: An Investigation of Personality Measures Associated with Over and Under Achievement in English and Arithmetic. 397
D. MARTYN: The Idea of God Among Roman Catholic Children Aged Seven and Eight 399
A. MCHALE: An Investigation of Personality Attributes of Stammering, Enuretic and School-phobic Children 400
R. A. NEWBERY: Objective Indices in the Assessment of Essays . . 403

CRITICAL NOTICES:

N. FRANCE: The Schools Council (1966). Examining at 16+. The Report of the Joint G.C.E./C.S.E. Committee of the Schools Council 406
G. KEIR: The i.t.a. Symposium 408

BOOK REVIEWS 414
PUBLICATIONS RECEIVED 431
INDEX TO VOLUME XXXVII 432
NOTES TO CONTRIBUTORS 436

8 FEB 1968